

UNIVERSITÉ DU QUÉBEC À MONTREAL

IDENTIFICATION ET CARACTÉRISATION MOLÉCULAIRE DES
PROTÉINES LIANT L'ARN LORS DE L'ACCLIMATATION AU
FROID CHEZ LE BLÉ (*TRITICUM AESTIVUM* L.)

MÉMOIRE
PRÉSENTÉ
COMME EXIGENCE PARTIELLE
À LA MAÎTRISE EN BIOLOGIE

PAR
MICHEL HECHEIMA

FÉVRIER 2006

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.01-2006). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

J'aimerais en premier lieu, remercier mon directeur de recherche le Dr. Fathey Sarhan qui m'a aimablement accueilli dans son merveilleux laboratoire, et de m'avoir ouvert la porte à ce passionnant domaine de recherche. Les mots me manquent pour vous témoigner mon énorme gratitude et vous remercier de votre générosité sans limite ainsi que votre optimiste contagieux.

Merci également aux professeurs Jean Danyluk, Mario Houde, François Dragon, Eric Rassart, et Normand Chevrier pour m'avoir aidé par vos conseils, corrections de travaux et aide technique lors de la réalisation de ce travail.

J'aimerais remercier spécialement Amadou Diallo, Jean-Benoît Frenette Charron, Mohammed Badawi, Nicolas Besnier, Ndjido Kane et Khalil Kane pour votre sincère amitié et de votre aide inestimable. Par votre belle humeur quotidienne, vous m'avez toujours encouragé et appuyé dans mes bons comme dans mes mauvais moments, et je vous en serai toujours reconnaissant. Nous avons eu un plaisir immense avec nos nombreuses blagues et potins qui agrémentaient nos journées au laboratoire. Vous avez rendu mon environnement de travail si agréable et parfois même délirant, ce qui constitue un excellent atout psychologique contre les déceptions occasionnelles de la recherche.

Une mention spéciale pour l'aide informatique et les différents échantillons biologiques prêtés par Mélanie Grondin. Tu as été aussi utile à notre projet qu'une chaperonne l'est pour une protéine. Merci également à Mahdi Belcaid et Mélanie Pelletier, pour votre aide bioinformatique et aussi pour votre grande patience. Un gros merci à Armand Miguel qui a participé à mon projet lors de son stage, et avec qui j'ai passé d'excellents moments.

J'aimerais remercier tous mes autres collègues du laboratoire pour leur coopération et leur professionnalisme. Les nombreuses discussions avec vous tous autour de la table de café et du gâteau hebdomadaire ont été très instructives et utiles.

J'aimerais faire mention de mes fidèles amis de toujours, Sania Mansouri, Ali Khalil Sébastien El Meknassi, Farizade Moulfî et Nicholas Svitek pour tous ce que vous avez fait pour moi et pour ma carrière scientifique. Sans vous, je ne serais pas ce que je suis aujourd'hui.

Cette page ne serait complète, sans mentionner l'apport illimité de ma famille qui m'a supporté et qui m'a tant donné d'énergie et de courage durant mes études. Ces remerciements sont spécialement dédiés à ma mère, Sonia Krikorian, pour tout son dévouement et son inconditionnel soutien à la réussite de ce projet. Tous les mots du monde ne sont pas assez pour te remercier du plus profond de mon cœur.

A vous tous merci, pour ces merveilleux souvenirs

TABLES DES MATIÈRES

LISTE DES FIGURES.....	vii
LISTE DES TABLEAUX.....	ix
LISTE DES ABRÉVIATIONS.....	x
RÉSUMÉ.....	xii
1.0 INTRODUCTION ET REVUE DE LITTÉRATURE.....	1
1.1 Introduction	1
1.2 Conditions climatiques	1
1.3 Tolérance au froid.....	2
1.4 Dommages causés par le gel.....	4
1.5 Modifications au niveau de la membrane plasmique.....	4
1.6 Génétique de l'acclimatation au froid chez le blé.....	5
1.7 Études sur la transmission du signal.....	7
1.8 Modifications au niveau des protéines	8
1.9 Les promoteurs	9
1.10 Mécanisme de contrôle de la transcription	10
1.11 Les facteurs de transcription.....	11
1.12 La répression	11
1.13 L'ARN polymérase II.....	12
1.14 La famille de transcription AP2.....	13
1.15 La sous famille CBF/DREB	14
1.16 Contrôle post-transcriptionnel de l'expression génique	15
1.17 Les différents rôles de la famille des protéines liant l'ARN.....	17
1.18 Le transport et la localisation de l'ARN	17
1.19 Recrutement et transport sur l'axe cytosquelettique.....	20
1.20 Stabilité relative des ARN chez les procaryotes et les eucaryotes.....	21
1.21 La structure de coiffe.....	21
1.22 Polyadénylation des ARNm	22
1.23 L'excision d'introns et ligation d'exons	22

1.24 L'épissage des ARN régule l'expression de protéines	23
1.25 Structures tridimensionnelles de l'ARN et de ses complexes.....	24
1.26 Le complexe ribonucléoprotéique hétérogène (hnRNPs)	25
1.27 Structures des motifs des protéines ayant la fonction de lier l'ARN	26
1.28 La stabilité des ARNm cytoplasmiques varie beaucoup d'un type à l'autre	28
1.29 La dégradation de certains messagers est sujette à une régulation	29
1.30 Les différents sous-groupes de protéines liant l'ARN.....	30

2.0 IDENTIFICATION ET CARACTÉRISATION MOLÉCULAIRES DES PROTÉINES LIANT L'ARN LORS DE L'ACCLIMATATION AU FROID

CHEZ LE BLÉ (*TRITICUM AESTIVUM* L.) 34

2.1 RÉSUMÉ.....	35
2.2 INTRODUCTION	36
2.3 MATÉRIELS ET MÉTHODES.....	40
2.3.1 Matériel végétal et conditions de croissance.....	40
2.3.2 Extraction d'ARN et de protéines totales	40
2.3.3 Analyses bioinformatiques	41
2.3.4 Northern Blot.....	41
2.3.5 Expression de la protéine recombinante	42
2.3.6 Purification de la protéine recombinante	43
2.3.7 Production d'anticorps	43
2.3.8 Western Blot.....	43
2.3.9 Test de résistance au froid des bactéries	44
2.3.10 Transcription in vitro	44
2.3.11 Northwestern blot	45
2.3.12 Test de liaison de l'ARN par retard sur gel	46
2.4 RÉSULTATS.....	47
2.4.1 Analyse moléculaire et structurale des séquences d'ADNc des protéines	47
2.4.2 Domaines auxiliaires retrouvés chez la famille des protéines liant l'ARN.....	50
2.4.3 Recherche des numéros d'accèsion d' <i>Arabidopsis thaliana</i> et de <i>Oryza sativa</i>	51
2.4.4 Expression du gène TaGr-rbp lors de l'acclimatation au froid et de stress.....	52
2.4.5 Production et purification de la protéine TaGR-RBP	53
2.4.6 Immunobuvardage de la protéine TaGR-RBP par l'anticorps polyclonal	53
2.4.7 Test de liaison.....	54

2.4.8 Étude phylogénique	54
2.5 DISCUSSION	55
2.6 RÉFÉRENCES	61
2.7 LÉGENDES DES FIGURES	63
2.8 FIGURES	68
3.0 CONCLUSION.....	85
4.0 RÉFÉRENCES.....	87
5.0 ANNEXE.....	95

LISTE DES FIGURES

Figure	Page
1 Séquence et organisation structurale des gènes codants des protéines liant l'ARN caractérisés chez le blé.....	68
2 Représentation schématique des différents motifs des gènes isolés ainsi que leurs longueurs en acides aminés.	72
3 Représentation schématique de la structure des protéines liant l'ARN de plusieurs organismes.	73
4 Accumulation des transcrits de <i>Tagr-rbp</i> durant l'acclimatation au froid chez différents génotypes de blé et d'espèces de céréales.....	74
5 Accumulation des transcrits de <i>Tagr-rbp</i> durant l'acclimatation au froid et lors de stress abiotiques.....	75
6 Accumulation des transcrits de <i>Tagr-rbp</i> durant l'acclimatation au froid chez le blé de printemps Manitou et le blé d'hiver Norstar.....	76
7 Localisation chromosomique de <i>TaGr-rbp</i> chez le blé (<i>Triticum aestivum</i> L)	77
8 Production de la protéine recombinante de TaGR-RBP chez la bactérie <i>E. Coli</i>	78

9	Purification de la protéine recombinante de fusion GST- TaGR-RBP produite par induction des bactéries <i>E. Coli</i> BL21 1A.....	79
10	Accumulation de la protéine TaGR-RBP chez le blé de printemps (Manitou) et d'hiver (Norstar) et les lignées isogéniques lors de l'acclimatation au froid.....	80
11	Test de liaison par northwestern blot du transcrit CBF1 radioactifs sur différents échantillons de protéines.....	81
12	Arbre phylogénique de gènes d' <i>Arabidopsis</i> et de blé réalisé par le "program paup 4.0 Beta".....	82

LISTE DES TABLEAUX

Tableau		Page
1	Liste de quelques domaines auxiliaires se trouvant chez la famille de protéines liant l'ARN.	95
2	Liste des numéros d'accèsion d' <i>Arabidopsis thaliana</i> ayant un domaine de liaison à l'ARN RRM et ceux de leurs homologues chez le riz avec leurs valeurs d'homologie (E value).	96

LISTE DES ABRÉVIATIONS

ABA	Acide abscissique
ADN	Acide désoxyribonucléique
ADNc	ADN complémentaire
ARN	Acide ribonucléique
ARNr	ARN ribosomiques
ARNm	ARN messenger
ARNt	ARN de transfert
ATP	Adénosine triphosphate
CPM	Coups par minute
Cs-RBD	Consensus sequence RNA-Binding domain
cv	Cultivar
DTT	Dithiothréitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Acide éthylènediamine tétraacétique
FGAS	Functional genomics of abiotic stress in crops
GST	Glutathione-S-Transferase
hnRNP	Ribonucléoprotéine hétérogène nucléaire
kDa	KiloDalton
mM	Millimolaire
MOPS	Acide 3-(N-morpholino) propanesulfonique
NCBI	National Center for Biotechnology information
pb	Paire de bases
PCR	Réaction de polymérisation en chaîne
RNase	Ribonucléase
RRM	Motif de reconnaissance d'ARN
SDS	Dodécylsulfate de sodium
SSC	Sodium citrate salin

Tris	Tris (hydroxyméthyl) aminométhane
Tween-20	Polyoxyéthylène sorbitan monolauréate
U	Unité
μ l	microlitre
U.V	Ultraviolet

RÉSUMÉ

Au cours de l'évolution, les espèces vivantes ont dû s'adapter à leur environnement externe afin de pouvoir vivre face aux différents stimuli. De la bactérie jusqu'à l'humain, les organismes qui avaient un avantage évolutif ont pu survivre et assurer la survie de leur espèce. Les plantes sont constamment soumises à des conditions de stress environnementaux qui mettent en péril leur croissance et leur développement. Les basses températures sont parmi ces stress et ils ont un impact économique important au Canada sur les cultures céréalières. Le blé s'acclimate au froid et différents cultivars ont une tolérance au gel variable. Les cultivars de blé d'hiver ont la capacité de résister à une baisse de température allant jusqu'à -27 °C contrairement aux cultivars de printemps. Les gènes modulés par le froid (COR) et leurs régulateurs CBF (Cis repeat Binding Factor) sont des gènes à la base de cette différence de tolérance. Les génotypes de blé qui accumulent le plus les transcrits de CBF1 montrent une plus grande tolérance au froid. Les bases moléculaires expliquant les différences d'expression de CBF sont inconnus. Cela peut être dû à des modifications post-transcriptionnelles ou une plus grande stabilité des transcrits. Cette stabilité peut être due aux protéines liant l'ARN. Ces dernières contrôlent aussi plusieurs aspects du métabolisme de l'ARN comme la maturation de ses extrémités, l'épissage, le transport, la localisation et l'efficacité de la traduction. Afin de comprendre le rôle de ces protéines dans la régulation de l'ARN cible, nous avons isolé et séquencé 4 gènes dont leurs protéines contiennent le motif RRM et des domaines auxiliaires riches en glycine. Plusieurs protéines contenant ces domaines auxiliaires sont régulés de façon circadienne tout comme les CBF. L'analyse d'expression démontre que le gène *TaGr-rbp* était exprimé au cours de l'acclimatation au froid, d'un choc thermique et de stress salin. Un anticorps polyclonal a été obtenu contre la protéine recombinante TaGR-RBP. L'immunobuvardage montre une accumulation jusqu'à 98 jours d'acclimatation chez les cultivars de blé Manitou et Norstar. Le gène a été localisé sur le chromosome 5 du génome du blé. Ce chromosome est reconnu comme porteur de plusieurs gènes impliqués lors de la régulation de la tolérance au gel. Des tests de liaison montrent une interaction entre la protéine TaGR-RBP et le transcrit de CBF1. L'analyse phylogénique a permis de classer les gènes isolés du blé dans le sous-groupe des protéines liant l'ARN riche en glycine. Des études de fonction pour surexprimer le gène *TaGr-rbp* chez les plantes transgéniques sont présentement en cours afin de déterminer la fonction.

Mots clés : blé, acclimatation au froid, stress abiotiques, protéine liant l'ARN, protéine riche en glycine

1.0 INTRODUCTION ET REVUE DE LITTÉRATURE

1.1 Introduction

Les plantes sont un élément essentiel à l'équilibre naturel sur terre. Elles produisent de l'oxygène et des sucres en utilisant le CO₂ et l'énergie solaire pour la photosynthèse. Par ce mécanisme important, les plantes jouent un rôle prépondérant dans le maintien de l'équilibre du CO₂ et permettent de contrer l'effet de serre sans cesse croissant. Elles sont à la base de la chaîne alimentaire en plus de servir dans divers domaines. L'industrie du bois pour la construction et l'immobilier n'est qu'un exemple montrant l'importance des plantes chez l'humain. Il ne faut pas oublier certaines plantes médicinales desquelles découlent plusieurs médicaments comme l'aspirine ou des drogues comme la morphine qui sert énormément comme un puissant anti-douleur dans le domaine médical (Boyle, 2001).

Au cours des cinq dernières décennies, la population mondiale a presque quadruplé alors que la production agricole mondiale n'augmentait que du quart. La production mondiale de plusieurs céréales telles que le riz, le blé, le maïs et l'orge est de loin inférieure à la demande (FAO, 2001). L'augmentation exponentielle de la population mondiale, provoquera d'importantes pénuries et favoriser ainsi l'augmentation des famines qui est déjà un fléau planétaire.

L'industrie agricole fait face à des défis importants pour augmenter la productivité et la qualité des céréales, mais également pour diminuer les pertes occasionnées par plusieurs facteurs dont les diverses maladies et les conditions climatiques.

1.2 Conditions climatiques

La température est aussi un des facteurs environnementaux qui influencent énormément la croissance et le développement des plantes. Dans les régions

nordiques, les chutes de températures soudaines ou l'arrivée d'un gel imprévu provoquent des dommages irréparables aux plantes d'intérêt agricole comme le blé, le maïs, la tomate, etc. Il est donc important de comprendre les mécanismes physiologiques et moléculaires impliqués dans la tolérance au froid.

Le caractère de résistance au froid a pu être manipulé et maximisé à l'intérieur de certaines espèces grâce aux croisements génétiques classiques mais l'épuisement des pools génétiques des espèces a rapidement mis un terme à l'amélioration de ce caractère. Il devient donc évident que dans l'avenir, les techniques traditionnelles de transfert de gènes par croisement ne seront pas suffisantes pour limiter les pertes ainsi que les nombreuses restrictions géographiques dues aux basses températures. Il devient donc impératif de développer de nouvelles approches biotechnologiques pouvant mener à une meilleure tolérance au froid des cultures et ainsi améliorer la productivité. Le développement des nouvelles technologies de génie génétique, permettant la transformation génétique des espèces par l'introduction de gènes étrangers, permet l'emploi de nouveaux programmes d'amélioration. Par contre, le transfert du caractère de tolérance au gel se heurte encore à des difficultés dues au manque de connaissance des mécanismes impliqués lors de la réponse au froid, des gènes impliqués ainsi que leur régulation (Danyluk, 1996).

1.3 Tolérance au froid

L'acclimatation au froid et la tolérance au gel impliquent un ensemble de modifications métaboliques, physiologiques et cellulaires (Thomashow, 1999; Pearce, 1999; Xin et Browse, 2000; Breton *et al*, 2000). Des modifications au niveau des lipides, de l'équilibre énergétique et hormonal ne sont que des exemples pour illustrer les différentes modifications métaboliques (Uemura *et al*, 1995; Perras et Sarhan, 1984; Dallaire *et al*, 1994). Il y a aussi l'augmentation du contenu en sucres, en proline, et en protéines solubles (Breton, 2003). Souvent, le rôle précis qu'auraient ces changements dans le processus d'acclimatation n'est pas très connu. Pour aborder

l'étude de la tolérance au gel, il faut identifier les modifications métaboliques et moléculaires induites au cours de l'acclimatation au froid, au moment où les cellules mettent en place les mécanismes, encore méconnus, conduisant au développement de la tolérance accrue au gel. La période d'acclimatation nécessaire pour atteindre la tolérance au gel maximale varie d'une espèce à l'autre. La tolérance au froid n'est pas un phénomène instantané mais plutôt induit lors d'une période d'exposition à des basses températures au dessus du point de congélation. Cette période est de l'ordre de 10 à 15 jours chez la pomme de terre, *Arabidopsis* et l'épinard, tandis qu'elle est environ de 50 jours chez le blé (Perras et Sarhan, 1984 ; Fowler *et al*, 1996). Au printemps, lorsque la température augmente, les plantes perdent la tolérance et reprennent leur croissance normale. Ce processus, appelé désacclimatation au froid, est rapide comparativement à l'acclimatation. Chez le blé, une exposition de cinq jours à une température de 20°C est suffisante pour qu'il y ait perte de tolérance au gel.

L'acclimatation au froid est associée à des modifications complexes au niveau de l'expression des gènes. Pour comprendre la nature multigénique de ces changements et les rôles des gènes spécifiques à l'acclimatation, il faut pouvoir isoler et caractériser ces gènes. Selon plusieurs études, l'expression des gènes des familles Wcs 120 (wheat cold specific), Wcor 410 et Wcor 518 (wheat cold regulated) démontre une corrélation positive avec la capacité des espèces à développer la tolérance au gel. Ces gènes sont exprimés à un niveau plus élevé chez le blé et le seigle d'hiver tolérant au gel, comparativement aux blés de printemps, à l'avoine et l'orge qui sont beaucoup moins tolérants. Les produits de ces gènes pourraient donc être impliqués dans le phénomène de la tolérance au gel. Les plantes tolérantes au froid sont capables de croître à des températures proches du point de congélation et de développer une tolérance au gel. Cette tolérance au gel permet aux plantes bisannuelles de survivre aux conditions rigoureuses de l'hiver et de reprendre leur croissance au printemps (Danyluk, 1996). Les mécanismes nécessaires à l'induction

de la tolérance au gel doivent être déclenchés au moment approprié et maintenus durant la période de risque. Cette capacité d'acclimatation est un atout majeur pour les espèces capables d'acclimatation puisque cela s'avère un caractère génétique d'un intérêt économique. Les homologues de tous les gènes isolés du blé sont présents dans les génomes de différentes graminées sensibles et tolérantes au gel. Par contre ces homologues ne sont pas exprimés chez les espèces sensibles tel le riz et le maïs.

1.4 Dommages causés par le gel

Ce dommage est associé à des changements de l'ultrastructure de la membrane plasmique durant la déshydratation induite par le gel. La conséquence de ces changements est que la membrane perd, dans certaines régions, sa structure de bicouche lipidique et par conséquent son intégrité biologique. Les membranes cellulaires sont reconnues depuis longtemps comme les principaux sites des dommages causés par le gel extracellulaire (Browse et Xin, 2001). Les différentes membranes cellulaires ne possèdent pas le même degré de sensibilité au gel et la membrane plasmique est considérée comme le site primaire de ces dommages. Les modifications biophysiques et ultra structurales de la membrane plasmique lors d'un gel extracellulaire ont été bien étudiées (Ouellet, 1998).

1.5 Modifications au niveau de la membrane plasmique

Durant l'acclimatation au froid, les contenus en sucres simples augmentent chez le blé (Perras et Sarhan, 1984). L'accumulation de composés tels le sucrose, le raffinose, le sorbitol, la glycine, la betaine et la proline a aussi été détectée chez plusieurs espèces durant divers stress osmotiques (Allard *et al*, 1998 ; Bohnert *et al*, 1995). Ils agiraient dans l'ajustement osmotique en aidant à retenir l'eau intracellulaire lors des stress hydrique et salin ou un stress provoqué par le gel. Ces composés peuvent stabiliser les membranes par une interaction avec les groupements polaires des phospholipides ou la formation d'interactions hydrophobes avec les membranes.

Les membranes cellulaires sont, depuis longtemps, reconnues comme les sites principaux de dommages causés par le gel. L'analyse des lipides de la membrane plasmique de plantes de seigle acclimatées et non-acclimatées indique que l'acclimatation au froid provoque une teneur plus élevée en stérols et en phospholipides (Steponkus et Lynch, 1989). La composition de la membrane plasmique a aussi été analysée chez *Arabidopsis* durant l'acclimatation au froid (Uemura *et al*, 1995). Cette analyse a révélé que l'acclimatation au froid provoque une teneur plus élevée en phospholipides et une teneur plus faible en stérols et en cérébrosides.

Une étude réalisée avec deux cultivars de blé possédant des capacités différentes d'acclimatation au froid démontre que certains changements structuraux sont associés au développement de la tolérance au gel (Zhou *et al*, 1994). Dans cette étude, l'augmentation des polypeptides a été spécifiquement détectée dans les membranes plasmiques isolées de feuilles d'un blé d'hiver. Ces auteurs ont remarqué des changements dans la structure de la membrane plasmique après l'acclimatation au froid. Ils notent que ce sont des protéines membranaires qui diminuent de façon plus marquée chez le blé d'hiver durant l'acclimatation au froid. Ces résultats montrent que, durant l'acclimatation au froid, la membrane plasmique est modifiée au niveau lipidique et protéique afin de mieux tolérer le gel (Danyluk *et al*, 1996).

1.6 Génétique de l'acclimatation au froid chez le blé

Le blé représente l'espèce la plus cultivée au niveau mondial, avec une production de 550 millions de tonnes au niveau mondial dont près de 27 millions de tonnes au Canada. Plus de 40% des aliments consommés par l'homme sont dérivés des grains de blé et de riz, alors que cinq autres céréales fournissent un autre 10%. La famille des graminées constitue donc la plus grande part de l'alimentation humaine. L'utilisation du blé comme modèle génétique pose toutefois certains problèmes. La

taille et la complexité de son génome en font une espèce difficile à manipuler et à analyser (Ouellet, 1998).

Le blé cultivé (*Triticum aestivum* L.) est une espèce hexaploïde composée de 3 génomes semblables mais non identiques (AABBDD, 42 chromosomes) provenant de parents diploïdes. Cette espèce a été obtenue suite au croisement de *Triticum tauschii* (génome diploïde D) avec le tétraploïde *Triticum turgidum* (génome AABB), ce dernier provenant du croisement entre *Triticum monococcum* (génome A) et un donneur du génome B, *Triticum searsii*. Il a été estimé que le génome du blé hexaploïde contenait 18.1 pg d'ADN par noyau haploïde, une quantité équivalente à environ $16-17 \times 10^9$ paires de bases. Plus de 75% est constitué de séquences répétitives présentes en quantité et en longueur différentes, alors qu'environ 20% consiste en séquences uniques ou présentes en faible nombre de copies. Les gènes codant pour des protéines ne représentent probablement qu'au plus 1% du matériel génétique nucléaire.

La taille du génome et la proportion de séquences répétées font que le blé est une espèce difficile à étudier au point de vue génétique. Heureusement, la disponibilité de nombreuses lignées (nullisomiques, télosomiques, ditélosomiques, substitution, etc.) produites à partir de la variété hexaploïde "Chinese Spring" facilite l'étude cytogénétique du génome de cette importante céréale. Ces individus aneuploïdes ont été utilisés pour localiser les gènes sur les chromosomes, pour cartographier les distances entre les gènes et les centromères, pour transférer des chromosomes d'une espèce à une autre, et pour identifier des homologies au niveau de chromosomes (McIntosh, 1987). La série ditélocentrique préparée dans "Chinese Spring" est composée de lignées où la paire de bras courts ou longs est manquante pour un chromosome en particulier. Ainsi, la lignée 6AL (L pour long arm) possède tous les chromosomes intacts sauf pour ce qui est de la paire de chromosomes 6A, où les bras

courts sont manquants. Les lignées ditélocentriques de blé sont très importantes puisqu'ils sont très utiles lors d'analyse de type Southern (Ouellet, 1998).

Ces résultats indiquent que les chromosomes du groupe 6 semblent porter plusieurs gènes structuraux, membres d'une famille apparentée, impliqués dans les réponses aux stress de basses températures et de déshydratation. La série de substitution 'Cheyenne'- "Chinese Spring" a permis de déterminer qu'un facteur positif important dans la régulation de l'expression de *wcs120*, *wcor410* et *wcor719* se trouvait sur le chromosome 5A (J. Danyluk, communication personnelle). Ce chromosome est celui où on a le plus fréquemment attribué une importance dans la tolérance au gel. Il est aussi celui sur lequel on retrouve un déterminant majeur de la tolérance, le locus *fr1* (frost resistance) (Galiba *et al*, 1995 ; Sutka, 1994). Chez l'orge, on trouve le plus important "quantitative trait loci" (QTL) pour la survie hivernale, la TL50, le phénotype de croissance et le contenu du collet en fructans, sur le bras long du chromosome 7 (équivalent au groupe 5 chez le blé) Hayes *et al*, (1993). Il semble donc que le chromosome 5A et ses homologues portent un ou plusieurs éléments qui contrôlent l'expression des gènes inductibles par les basses températures et qui sont corrélés avec la tolérance au gel.

1.7 Études sur la transmission du signal

Pour induire le phénomène d'acclimatation au froid, menant à la tolérance au gel, la plante doit posséder un système capable de percevoir la température ambiante et ensuite de déclencher les processus biochimiques impliqués. Peu d'études se sont penchées précisément sur cet aspect. Durant une exposition au froid de trois heures, la phosphorylation d'une protéine était augmentée de 12 fois (Monroy *et al*, 1993). De plus, le développement de la tolérance peut être aboli par l'utilisation de bloqueurs de canaux de calcium et par l'antagoniste des protéines liant le calcium telle la calmoduline. Ces résultats suggèrent que le calcium serait impliqué dans la transmission du signal. Dans une étude réalisée par Monroy et Dhinsa (1995), l'influx

de $^{45}\text{Ca}^{2+}$ à 4°C chez les protoplastes de luzerne était 15 fois plus élevé qu'à 25°C. L'addition de chélateurs de calcium ou de bloqueurs de canaux de calcium inhibe l'influx de calcium extracellulaire et l'expression des gènes *cas* à 4°C. Par contre, l'addition d'un ionophore de calcium ou un agoniste des canaux de calcium à des protoplastes non acclimatés cause l'influx de calcium extracellulaire et l'induction des gènes *cas* à 25°C. L'augmentation était transitoire, suggérant que l'influx de calcium peut induire mais non maintenir l'expression des gènes *cas*. Les auteurs concluent que le calcium jouerait un rôle essentiel dans la transmission du signal et que des événements encore non élucidés tels que l'induction de protéines kinases pourrait provoquer une expression différentielle de gènes durant l'acclimatation au froid et ainsi le développement de la tolérance au gel.

1.8 Modifications au niveau des protéines

Plusieurs études, résumés par Singh et Laroche (1988) et Guy (1990) ont rapporté que l'acclimatation au froid provoquait une augmentation des contenus en ARN totaux et en protéines solubles. De plus, l'ajout de cycloheximide, un inhibiteur de la synthèse protéique, inhibe le développement de la tolérance au gel. Ceci a permis de proposer que l'augmentation de la tolérance au gel chez les plantes, durant l'acclimatation au froid, était le résultat d'une combinaison de changements qui dépendent de la modification de l'expression génétique. Les travaux effectués par Sarhan et Chevrier (1985) ont démontré que l'augmentation de la synthèse d'ARN était due à une augmentation marquée de l'activité de l'ARN polymérase I et à un plus faible degré de l'ARN polymérase II chez un blé d'hiver. Cette augmentation n'était pas seulement due à la croissance aux basses températures car elle était moins marquée chez un blé de printemps dans les mêmes conditions de croissance. L'augmentation de l'activité de l'ARN polymérase II était la première évidence qu'il y avait une augmentation de transcription de gènes lors du processus d'acclimatation.

Les analyses électrophorétiques des protéines de plantes non acclimatées et acclimatées au froid ont révélé plusieurs changements qualitatifs et quantitatifs. Ceci démontrait que les plantes tolérantes au gel avaient de nouvelles protéines qui n'étaient pas présentes chez les plantes non acclimatées. Ce type d'étude ne pouvait pas informer sur la fonction des différentes protéines exprimées différenciellement durant l'acquisition de la tolérance au gel (Castonguay *et al*, 1993). C'est pourquoi après l'identification des protéines, des études plus approfondies sont nécessaires pour déterminer les possibles fonctions des protéines différenciellement exprimées.

1.9 Les promoteurs

Dans le but d'identifier les éléments qui confèrent la spécificité d'expression durant l'acclimatation au froid, plusieurs équipes ont entrepris l'étude des régions promotrices des gènes régulés par le froid. Pour le moment, l'identification des éléments impliqués dans la régulation des gènes induits durant l'acclimatation au froid n'est que partielle. Une caractérisation plus détaillée des promoteurs de gènes régulés par le froid permettra d'identifier les facteurs cis et trans qui sont responsables de la régulation par le froid, de la spécificité tissulaire et cellulaire et du contrôle de l'expression associée au développement. L'information apportée par ces études aidera à mieux diriger une surexpression des gènes seulement chez les cellules et tissus qui nécessitent ces gènes pour le développement de la tolérance au gel (Danyluk, 1996).

Il existe des centaines de promoteurs dans le génome des cellules procaryotes et des milliers dans celui des cellules eucaryotes. La fréquence à laquelle la transcription démarre sur un gène est généralement en rapport avec le degré d'utilisation du produit de ce gène. Par exemple, dans les cellules en division active, les gènes des ARN ribosomiques sont transcrits à fréquence élevée (Horton, 1994).

1.10 Mécanisme de contrôle de la transcription

Tant chez les procaryotes que chez les eucaryotes, il se forme un complexe de transcription efficace dès que la séquence d'ADN promotrice est repérée par une ou plusieurs protéines qui se fixent sur le promoteur, ainsi que sur l'ARN polymérase. Ces protéines appelées facteurs de transcription sont nécessaires à l'assemblage de l'ARN polymérase au site d'initiation. La séquence nucléotidique d'un promoteur est l'un des plus importants facteurs déterminant la vitesse à laquelle un gène est transcrit. L'étude de la structure de divers promoteurs est devenue possible dès la mise au point des techniques de séquençage de l'ADN. Des études ont établi des séquences consensus, qui se trouvaient en amont du site d'initiation de régions promotrices de différents gènes (Williams *et al*, 1992 ; Bray, 1993). Dans plusieurs cas, des séquences promotrices ne possèdent aucune concordance avec la séquence consensus. Ce sont des promoteurs peu puissants et font généralement partie des gènes transcrits à faible fréquence. D'autres séquences promotrices, comme celle des opérons d'ARNr chez *E. coli* correspondent de près à la séquence consensus. De nombreuses études ont montré que les gènes qui possédaient une séquence promotrice consensus sont toujours transcrits avec une grande efficacité. Ceci illustre que cette séquence est la plus efficace pour l'holoenzyme d'ARN polymérase. La sélection naturelle aurait modelé la séquence promotrice de chaque gène de façon à répondre aux besoins de la cellule. Un promoteur peu efficace résulte en une faible transcription du gène correspondant (Horton, 1994).

On connaît des gènes exprimés très activement dans certaines circonstances et complètement silencieux à d'autres moments. Ce sont des gènes soumis à une régulation contrôlée. Le démarrage de la transcription dépend à la fois de séquences spécifiques et de protéines solubles (facteurs de transcription).

1.11 Les facteurs de transcription

Les facteurs de transcription sont généralement composés de trois domaines fonctionnels, un domaine de liaison à l'ADN requis pour la reconnaissance et la fixation à l'élément de régulation, un domaine de multimérisation requis pour la formation d'homo- ou hétéromultimères et un domaine d'activation requis pour stimuler la transcription (Tijan et Maniatis, 1994). Les facteurs de transcription sont classés en familles selon la structure de leur domaine de liaison à l'ADN ou de multimérisation (Pabo et Sauer, 1992 ; Luscombe *et al*, 2000). L'expression des gènes requiert un contrôle serré de l'activité des facteurs de transcription. Pour ce faire, il existe plusieurs mécanismes qui contrôlent non seulement l'activité mais aussi la quantité des facteurs de transcription. La quantité de facteurs de transcription peut être régulée à plusieurs niveaux, c'est-à-dire au niveau de la transcription, de la maturation et stabilité des ARN messagers ou de la traduction sans oublier la protéolyse (Calkhoven et Ab, 1996). Une fois le facteur synthétisé, son activité est souvent contrôlée par différents mécanismes. La phosphorylation est la modification post-traductionnelle des facteurs de transcription la mieux caractérisée. Elle peut influencer la localisation cellulaire du facteur, sa capacité à lier l'ADN ou encore ses interactions avec d'autres composantes de la machinerie transcriptionnelle (Hunter et Karin, 1992). D'autres facteurs de transcription doivent recruter un facteur accessoire pour pouvoir communiquer avec l'ARN polymérase.

1.12 La répression

Il existe plusieurs exemples de répression transcriptionnelle, il est possible d'élaborer différents modèles pour la régulation négative de la transcription. Le modèle le plus simple de répression de la transcription est lorsqu'un répresseur entre en compétition avec un activateur pour la liaison à un site unique ou a des sites différents qui se superposent. Un autre modèle de répression implique la neutralisation d'un activateur. Dans ce cas, le répresseur n'empêche pas la liaison de

l'activateur à l'ADN mais inhibe sa capacité d'activer la transcription en bloquant les contacts avec l'ARN polymérase (holoenzyme).

D'autres répresseurs ne bloquent pas toujours l'action d'un activateur mais interfèrent plutôt avec la machinerie transcriptionnelle de base. Un mécanisme similaire au modèle précédant est utilisé par certains répresseurs pour bloquer cette fois-ci la formation du complexe d'initiation minimal. Pour ce type de répression, l'élément lié par le répresseur doit être à proximité de la boîte TATA (Boyle, 2001). Finalement, un autre modèle illustre la répression active où le répresseur, par un mécanisme similaire à l'activation de la transcription, agit sur l'ARN polymérase pour altérer son activité via un domaine de répression spécifique. Les domaines de répression active sont très divergents et peuvent être regroupés en trois catégories : les domaines riches en alanine, ceux riches en proline et les domaines chargés (Hanna-Rose et Hansen, 1996). Un domaine de répression active unique aux plantes vient tout juste d'être identifié (Ohta *et al*, 2001). Ce domaine contient le motif (L/F) DLN (L/F) XP, qui est présent chez plusieurs répresseurs de la transcription de plantes (Boyle, 2001).

1.13 L'ARN polymérase II

Les mécanismes généraux de la transcription par l'ARN polymérase II sont conservés au cours de l'évolution et les principes généraux de la transcription par cette enzyme sont applicables tant chez les eucaryotes unicellulaires que chez les organismes multicellulaires. La transcription dépendante de l'ARN polymérase II requiert l'assemblage de plusieurs protéines au promoteur. L'initiation de la transcription requiert la reconnaissance de la boîte TATA par le facteur TFIID. Ce dernier contient la protéine liant la boîte TATA (TBP) et ses facteurs associés. TFIIB interagit par la suite avec le facteur TFIID lié à la boîte TATA. Il est aussi responsable d'attirer l'ARN polymérase au complexe d'initiation minimal. Suite à l'action combinée de TFIID, TFIIB, et TFIIF, la polymérase est maintenant

positionnée au promoteur et peut dans des conditions précises, initier la transcription (Parvin et Sharp, 1993). Le recrutement de TFIIE permet l'ajout subséquent de TFIIH pour former le complexe d'initiation complet. Ce dernier est capable de toutes les fonctions associées à l'ARN polymérase II, c'est-à-dire l'élongation d'une chaîne d'ARN sur un modèle d'ADN par l'entremise d'une activité hélicase ADN dépendante de l'hydrolyse de l'ATP. L'expression d'un gène est sous le contrôle de plusieurs facteurs dont la topologie de l'ADN, les activateurs et les répresseurs de la transcription.

1.14 La famille de transcription AP2

L'incapacité de certaines espèces à développer une tolérance peut donc être due à une voie de transmission de signal inefficace et/ou à des éléments cis/trans inadéquats pour favoriser la transcription de gènes en ARNm. La régulation de la transcription est en grande partie modulée par des protéines liant l'ADN à des sites spécifiques situés dans les régions promotrices et activatrices des gènes. La plupart des études portant sur l'expression des gènes inductibles par les basses températures ont donc été orientées au niveau de la régulation de leur transcription (Meshi et Iwabuchi, 1995). Les études fonctionnelles sont basées sur la capacité des promoteurs à favoriser l'expression d'un gène rapporteur à basses températures, soit dans des plantes transgéniques ou dans des tests d'expression transitoires effectués par bombardement de tissus à l'aide de micro projectiles (Ouellet, 1998). Peu d'éléments cis impliqués spécifiquement dans la réponse aux basses températures (LTRE, "low temperature-responsive element") ont été identifiés jusqu'à maintenant. Plusieurs promoteurs possiblement inductibles par les basses températures possèdent des boîtes G (CACGTG), tout comme plusieurs gènes contrôlés par des facteurs environnementaux (Williams *et al*, 1992). Cet élément est le noyau de plusieurs éléments de régulation cis, incluant les éléments de réponse à l'ABA (ABRE) dont le consensus est RYACGTGGYR (Ingram et Bartels, 1996).

Il existe plusieurs gènes qui sont spécifiquement induits lors de basses températures comme par exemple Wcs 200 (Ouellet *et al*, 1993) et Wsc 120 (Gray *et al*, 1997). D'autres répondent à différents stress biotiques ou abiotiques comme la famille de gènes AP2, qui codent pour des protéines de régulation (Danyluk *et al*, 1998). Les membres de cette famille contiennent 1 ou 2 domaines conservés AP2. Ils sont très souvent retrouvés chez les plantes, mais la présence d'homologues a été retrouvée aussi chez les cyanobactéries et des virus (Magnani *et al*, 2004)

Le domaine AP2 est un motif d'environ 70 acides aminés présent dans la protéine et il est composé de 2 éléments YRG et RAYD (Jofuku *et al*, 1994 ; Riechmann et Meyerowitz, 1998). L'élément YRG se retrouve dans la région N-terminale qui est riche en résidus basiques et hydrophiles. L'élément RAYD se retrouve du côté C-terminal et est capable de former une hélice amphipatique alpha. La liaison ADN-AP2 semble dépendre de l'élément YRG.

1.15 La sous famille CBF/DREB

Les gènes de cette sous-famille sont induits par un ou plusieurs stress environnementaux et les fonctions sont directement reliés à la tolérance au gel et aux stress abiotiques chez les plantes. On en retrouve 56 et 58 membres chez *Arabidopsis thaliana* et le riz respectivement.

Stockinger *et al*, (1997) ont isolé un gène codant pour une protéine, CBF1, qui possède la capacité de lier in vitro le motif CCGAC de *Arabidopsis* et ce, spécifiquement. Comme ce motif a été impliqué dans la réponse à la sécheresse et aux basses températures, CBF1 doit être sérieusement considérée comme un facteur de transcription pouvant favoriser la transcription de gènes inductibles par les basses températures. La protéine encodée par ce gène, CBF1 (Cis repeat Binding Factor), a un poids moléculaire de 24 kDa et contient un domaine conservé liant l'ADN (AP2). Deux homologues additionnels de CBF1 (CBF2 et CBF3) dont les messages sont

régulés par les basses températures ont été isolés (Gilmour *et al*, 1998). Ces deux facteurs se lient spécifiquement au motif DRE et sont capables comme leur homologue CBF1 d'activer la transcription aussi bien dans les protoplastes foliaires que chez la levure.

Des expériences avec des plantes transgéniques supportent le rôle joué par ces potentiels facteurs de transcription dans la régulation de gènes induits par le froid et la sécheresse. Des plantes transgéniques surexprimant CBF1/DREB1B ont été produites en plaçant l'ADNc *cbf1* sous le contrôle du promoteur 35S (Jaglo-Ottosen *et al*, 1998). Les résultats montrent que la surexpression de CBF1 induit l'expression des gènes cor (cold regulated) chez les plantes transgéniques d'*Arabidopsis* non acclimatées au froid et par conséquent non stimulées par les basses températures. La tolérance au gel de ces plantes s'en trouve nettement améliorée. Des résultats similaires ont été observés chez des plantes transgéniques surexprimant constitutivement le facteur CBF3/DREB1A (Liu *et al*, 1998; Gilmour *et al*, 2000). En résumé, CBF1 et CBF3 sont responsables de l'accumulation des protéines COR, de l'accumulation de proline et de sucres solubles. Ces résultats montrent une grande implication de ces facteurs de transcription lors de l'acclimatation au froid et de la tolérance au gel.

1.16 Contrôle post-transcriptionnel de l'expression génique

La plupart des études de l'expression génique portent sur la régulation transcriptionnelle. Toutefois, la régulation peut aussi s'effectuer au niveau post-transcriptionnel comme l'épissage de l'ARNm. Elle peut être également au niveau sa stabilité relative, son efficacité traductionnelle, modification post traductionnelle et au niveau du catabolisme des protéines. Les études de transcription active de noyaux effectuées chez le blé, l'orge, *Arabidopsis* et la luzerne ont démontré que la régulation de l'expression de certains gènes inductibles par le froid pouvait

s'effectuer au niveau post-transcriptionnel (Chauvin, 1993 ; Dunn *et al*, 1994 ; Hajela *et al*, 1990 ; Wolfrain *et al*, 1993)

La régulation de l'expression au niveau post-transcriptionnel pourrait s'effectuer au niveau de la stabilité des messagers (Abler et Green, 1996). Il est possible que les protéines liant l'ARN puissent jouer un rôle dans la stabilisation sélective des ARNm, diminuant ainsi le taux catabolique (turnover) de ces messagers. Ceci pourrait permettre le maintien de l'efficacité de traduction afin de produire des niveaux adéquats de protéines de tolérance au gel. Malheureusement, peu d'études ont porté sur cet aspect de la régulation des gènes inductibles par le froid. En fait, les études publiées à ce jour portent sur l'analyse de la stabilité des messagers de gènes codant pour des protéines antigel de poisson (AFP, antifreeze protein) et de plants de tabac transgéniques pour un gène *afp*. Ces études ont révélés que la demi-vie des messagers *afp* était plus longue à basses températures qu'à la température de la pièce (Duncker *et al*, 1995). L'homologie des protéines AFP avec des protéines induites par le froid chez les plantes suggère la possibilité qu'une stabilisation des messagers à basses températures pourrait favoriser l'accumulation de protéines responsables de la tolérance au gel.

Il est connu que les mécanismes qui régulent la traduction ou la stabilité des messagers sont fondamentalement plus rapides à survenir que les mécanismes transcriptionnels, qui impliquent de multiples événements (Gallie, 1993). Cela suggère que les modifications post-transcriptionnelles pourraient être employées préférentiellement dans le cas des réponses qui doivent être flexibles et rapides, comme c'est le cas pour la réponse aux basses températures. Il s'avère donc nécessaire dans le futur d'étudier la régulation des gènes inductibles par le froid tant au niveau transcriptionnel que post-transcriptionnel.

1.17 Les différents rôles de la famille des protéines liant l'ARN

L'ARNm eucaryote n'est pas seulement une séquence linéaire codant pour une protéine. Chaque molécule d'ARN messenger possède un lot d'information importantes conditionnant sa future post transcription. Cela implique son transport du noyau, sa localisation, sa traduction, sa stabilité et sa dégradation. Il est envisageable que chacune de ces informations soit spécifiée indépendamment par le biais de séquences consensus placées sur l'ARN en des positions conservées. Néanmoins de plus en plus de résultats indiquent que ces fonctions ne seraient pas réalisées par l'ARNm lui-même mais par des protéines de liaison à l'ARNm assemblées en un complexe se fixant sur l'ARNm en un site précis. Les pré-ARNms ainsi que les ARNms ne sont jamais au cours de leur vie des acides nucléiques isolés. Ils existent sous forme de particules ARN-protéines (RNPs) (Dreyfuss *et al*, 1993). Le contenu protéique de ces particules évolue tout au long de leur existence, du début de la transcription jusqu'à l'exportation cytoplasmique et de la traduction jusqu'à la dégradation de l'ARN. Un des changements est le remplacement du CBP (complexe de liaison à la coiffe des transcrits de Pol II) par eIF4e dès l'exportation cytoplasmique des messagers. On sait également que certaines particules ribonucléiques hétérogènes (hnRNPs) se lient au pré-ARNm dès leur synthèse et les accompagnent dans le cytoplasme, où elles peuvent alors réguler les événements subis par l'ARNm. Plusieurs recherches démontrent aussi que l'épissage peut affecter la composition protéique des complexes protéines-ARN. Une étude a déjà prouvée que l'épissage augmente drastiquement l'efficacité d'expression de certains gènes (Luo et Reed, 1999).

1.18 Le transport et la localisation de l'ARN

L'établissement de la polarité cellulaire pendant le développement est régi par une répartition inégale de déterminants qui permettent les spécialisations fonctionnelles et structurales de ses compartiments. Les étapes de la différenciation cellulaire lors du développement nécessitent une coordination parfaite dans ses événements. Même

après le développement, plusieurs conditions qui exigent une adaptation en réponse à un stimulus interne ou externe nécessitent une répartition asymétrique de l'outillage biochimique dans la cellule (Duchaine, 2001).

Pourquoi alors transporter de l'ARN plutôt que de favoriser celui des protéines ? De multiples stratégies sont efficacement utilisées par la cellule pour spécialiser la biochimie de ses divers compartiments. Plusieurs stratégies dépendent du transport des protéines. L'asymétrie cellulaire peut être établie par une répartition différente des protéines impliquées dans des fonctions similaires. Cette ségrégation est aussi assortie d'un contrôle strict du trafic entre les compartiments.

Premièrement, cette stratégie est économe énergétiquement parlant. Le transport d'un transcrit permet la production locale d'une multitude de copies de la protéine encodée, amplifiant ainsi l'activité biochimique en un endroit défini pour un coût énergétique moindre en transport. Le deuxième avantage est que le transport d'ARN est très efficace pour limiter l'expression de l'activité d'une protéine. Cet avantage provient du couplage du transport d'un transcrit avec le contrôle de sa traduction (Duchaine, 2001). Un transcrit transporté n'est traduit qu'au site précis de sa localisation cellulaire. Ainsi, puisque les ARNm ne libèrent l'activité des protéines qu'ils encodent qu'au moment de la traduction, le transport de l'ARN a pour conséquence d'en limiter l'activité en un espace et un temps définis par la traduction locale du transcrit. Un exemple de cet avantage est le transport du transcrit de la protéine basique de la myéline (MBP) dans l'oligodendrocyte (Carson *et al*, 2001). Cette protéine permet la compaction des feuilletts de myéline de la cellule gliale autour de l'axone d'un neurone, optimisant de ce fait la rapidité et l'efficacité de la transmission des vagues de dépolarisation. Cette protéine se lie si fortement aux membranes, qu'il serait excessivement exigeant de la transporter jusqu'aux sites où sa fonction est désirable. Il devient donc impératif d'en limiter la traduction aux endroits où elle entrera en contact avec la membrane.

Le contrôle traductionnel d'un transcrit transporté implique donc, en principe, des signaux en cis et des déterminants en trans qui permettent, d'une part une répression du transcrit pendant le transport et d'autre part, une activation au moment et à l'endroit où l'expression est essentielle à la fonction. Bien que la nécessité de ces événements est en principe évidente pour réaliser l'expression localisée de l'ARNm, peu d'éléments et de facteurs impliqués ont été identifiés. Pour cette fonction, Staufén chez la drosophile demeure probablement la protéine la mieux caractérisée (Meyer et Gavis, 2005 ; Gautrey *et al*, 2005)

Un autre avantage du transport des transcrits est que le mécanisme offre une multitude de cibles de régulation qui s'ajoutent à celles qui sont disponibles lors du transport de protéines. Que ce soit au niveau de la ségrégation du transcrit, de sa stabilisation ou à l'une ou l'autre des étapes de la traduction, les facteurs impliqués offrent autant de cibles supplémentaires qui permettent une régulation fine de l'expression génétique. La précision de cette régulation est particulièrement importante dans un système où une différence minime entre des concentrations protéiques précises d'activité peut générer des réponses biologiques distinctes. Il est présentement assumé que le transport d'ARN débute par la reconnaissance de signaux de transport dans un transcrit par une ou plusieurs protéines liant l'ARN. Un peu comme pour les signaux de ciblage des protéines membranaires entre les compartiments cellulaires, des « adresses postales » constituées d'une combinaison de structure et de séquence comme les tiges boucles sont encodées dans les transcrits transportés. Ces signaux, presque exclusivement retrouvés en 3'UTR, dictent la voie à emprunter et la destination finale par le type de protéine(s) qu'ils recrutent. Bien que cet assemblage puisse se faire dans le cytoplasme, pour plusieurs transcrits il semble qu'une ou des étapes critiques de cette reconnaissance se produit dans le noyau (Duchaine, 2001).

Comme les pré-ARNm nucléaires et les ARNm matures prêts au transport au cytoplasme, les ARNm cytoplasmiques sont associés à des protéines pour former des complexes ribonucléoprotéiques. Le rapport protéine - ARN est cependant plus faible dans les RNPm cytoplasmiques que dans celles du noyau. La protéine de RNPm cytoplasmique prépondérante est la protéine cytoplasmique se liant au poly-A (PABP), fermement attachée à la queue poly (A) des ARNm.

1.19 Recrutement et transport sur l'axe cytosquelettique

La reconnaissance du transcrit par les protéines agissant en trans cause la formation d'un complexe ribonucléoprotéique (RNP). Or, une caractéristique très répandue pour ce type de complexe est leur très grande taille. Les interactions à l'origine de ces immenses complexes sont peu connues. Il est présentement proposé que la dimérisation ou la multimérisation de plusieurs des facteurs agissant en trans puisse contribuer à la formation de complexes d'une telle envergure (Duchaine, 2001). Pour distribuer de façon asymétrique les transcrits dans la cellule, il doit y avoir des structures ou des "pistes" qui en relient les différents compartiments. Pendant le transport de l'ARN, ce rôle est joué par le cytosquelette. Cette fantastique composante cellulaire, à la plasticité remarquable, assure plusieurs fonctions critiques dans le mécanisme de transport d'ARN.

Le modèle suggéré propose donc que le RNP assemblé soit associé avec un moteur permettant de fournir la force motrice pour le mouvement le long de l'axe cytosquelettique. Pour plusieurs transcrits, ce recrutement se fait en plusieurs étapes et peut même impliquer des composantes distinctes de cytosquelette. Le transport d'un transcrit implique, en sa dernière étape, son ancrage. Encore une fois à ce niveau, notre modèle implique une contribution du cytosquelette. Or, les interactions qui régissent l'ancrage peuvent être complètement distinctes de celles réalisant le transport actif du transcrit (Duchaine 2001).

1.20 Stabilité relative des ARN chez les procaryotes et les eucaryotes

Les molécules d'ARN ribosomiques (ARNr) sont intégrées aux ribosomes et représentent environ 80% de tout l'ARN cellulaire. Les ARNm qui sont reconnus pour être les acides nucléiques les moins stables ne représentent que 3 % de l'ARN cellulaire total. Ce type d'ARN est dégradé peu de temps après avoir été transcrit, contrairement à l'ARNr et l'ARNt qui sont des molécules très stables. Dans les cellules bactériennes, la moitié des ARNm sont dégradés par des nucléases dans les 3 minutes suivant leur transcription. Chez les eucaryotes, la demi vie moyenne des ARNm est d'environ 10 heures. Cette stabilité relativement élevée comparativement aux cellules procaryotes provient de sa maturation et d'autres modifications qui empêchent sa dégradation au cours de son passage du noyau au cytoplasme (Horton, 1994).

1.21 La structure de coiffe

Toutes les molécules d'ARNm eucaryotes subissent des transformations qui les rendent plus stables et en font de meilleurs substrats pour la traduction. Une façon d'accroître la stabilité est de modifier les extrémités de la chaîne en vue de les rendre insensibles aux exonucléases cellulaires qui dégradent l'ARN. L'extrémité 5' des ARNm eucaryotes se transforme avant que leurs précurseurs aient été entièrement synthétisés. L'extrémité 5' du transcrit primaire est un résidu de nucléoside triphosphate correspondant au premier nucléotide incorporé par l'ARN polymérase II. Cette extrémité est d'abord remaniée par l'élimination du groupe phosphate terminal sous l'action d'une phosphohydrolase (Horton *et al*, 1994) Le groupe 5'-diphosphate formé réagit alors avec le phosphore alpha d'une molécule de GTP pour donner une liaison 5'-5' triphosphate. Cette réaction est catalysée par une enzyme appelé guanylyl transférase et la structure résultante est une coiffe. Cette structure est caractérisée par la présence d'un résidu guanosine méthylé en position N7 relié par une liaison caractéristique 5'-5' au premier nucléotide de l'ARNm (Bisaillon, 2001). La coiffe transforme le précurseur d'ARNm en substrat pour d'autres enzymes

nucléaires de maturation comme ceux qui effectuent l'excision-épissage. Dans l'ARNm définitif, la coiffe sert à l'ancrage sur les ribosomes en vue de la synthèse protéique.

1.22 Polyadénylation des ARNm

L'ARNm est aussi modifié à son extrémité 3'. Dès que la polymérase II a transcrit au-delà du bout 3' de la séquence codante de l'ADN, l'ARN naissant est scindé au voisinage d'un site dont la séquence consensus est AAUAAA. La scission se produit à une distance de 10 à 20 nucléotides en aval de cette séquence et est appelé polyadénylation. Au bout 3' du transcrit primaire, le remaniement consiste en une coupure endonucléasique qui produit un groupe hydroxyle 3' libre auquel viendra s'attacher une série de résidus adénylate transférés par un enzyme appelé poly-A polymérase ; la queue poly-A formée comprend, selon l'espèce, de 100 à 250 nucléotides. Cet enzyme fonctionne sans amorce et ajuste de lui-même le nombre à peu près correct de résidus A nécessaires (Horton *et al*, 1994). Après la scission au site de polyadénylation, la poly-adénylation procède en deux temps : les 12 premiers nucléotides A sont attachés lentement, puis les 200-250 suivants très rapidement. Cette phase rapide exige la fixation de nombreux exemplaires d'une protéine à motif RNP, très avide de poly-A : on la désigne PABII, pour la distinguer de la première protéine de fixation au poly-A, celle qui se fixe à la queue poly-A des ARNm cytoplasmiques. Le complexe ARN-protéine sur la queue poly A stabilise l'ARNm en le protégeant d'une dégradation de l'extrémité 3' (Lodish *et al*, 1993).

1.23 L'excision d'introns et ligation d'exons

Les sites d'épissage des pré-ARNm comportent de courtes séquences universelles. On a trouvé la place des sites d'épissage des pré-ARNm en comparant la séquence des ADN génomiques à celle des ADNc préparés à partir des ARNm correspondants. Les séquences présentes dans l'ADN génomique et absentes de l'ADNc représentent les introns et marquent l'emplacement des jonctions intro-exon. L'étude d'une grande

série d'ARNm a fait apparaître de courtes séquences consensus modérément conservées aux jonctions intro-exon des pré-ARNm eucaryotes et les organismes supérieurs. Les seuls nucléotides universellement conservés sont GU en 5' et AG en 3' de l'intron. L'étude des délétions introduites dans la portion centrale des introns de différents pré-ARNm a montré que 30 à 40 nucléotides à chaque extrémité d'un intron suffisent en général à assurer la vitesse normale d'excision et d'épissage.

L'utilisation d'extraits nucléaires capables d'exciser les introns d'une chaîne d'ARN et d'en unir les exons aux sites corrects fut la clé qui révéla le mécanisme de l'épissage de l'ARN qui est primordial pour la régulation des gènes. L'intron s'élimine sous forme de boucle appelée lasso, une structure dans laquelle le G du bout 5' de l'intron est lié par une liaison phosphodiester 2'-5'. Cette liaison est inhabituelle à un résidu adenosine proche du 3' de l'intron. Cette adenosine sert de point d'embranchement, car c'est à cet endroit que la chaîne d'ARN se referme sur elle-même en boucle. Le fait que les introns sont excisés sous forme de structure ramifiée en boucle révéla que l'excision et l'épissage passe par 2 réactions de trans-estérification. Le bilan net de ces deux réactions consiste à la réunion de 2 exons et l'élimination de l'intron intercalé sous forme de structure ramifiée. L'analogie entre des protéines végétales induites par le froid et les hnRNP animales suggère que les plantes pourraient posséder la capacité de faire de l'épissage alternatif des messagers en réponse à certains stress ou autres signaux environnementaux (Hughes et Dunn, 1996 ; Albà, et Pagès, 1998).

1.24 L'épissage des ARN régule l'expression de protéines

Un des exemples les mieux connus est le transcrit de fibronectine. Ce transcrit comporte plusieurs exons et l'épissage alternatif produit une variété de protéines. Ces protéines sont en général secrétées par les fibroblastes et les hépatocytes. Le transcrit primaire de fibronectine comporte de nombreux exons ; le choix selon lequel ils sont épissés produit toute une variété de protéines, secrétées par les fibroblastes et les

hépatocytes (aussi par d'autres types cellulaires). Les protéines des fibroblastes est une forme sécrétée de protéine fibreuse, qui s'accroche aux composants de la matrice extracellulaire et aux récepteurs de la surface des cellules, ancrant par ce mécanisme les cellules à la matrice extracellulaire. Dans le fibroblaste, l'épissage du pré-ARNm de fibronectine produit un ARNm de deux exons, EIIIA et EIIIB, qui codent des domaines protéiques d'interaction avec les récepteurs de surface de maints types cellulaires. La fibronectine de fibroblaste ira donc couvrir la surface d'un grand nombre de cellules. Par contre, l'épissage de la fibronectine d'hépatocyte "saute" ces deux exons, pour donner une protéine Tra, dépourvue de domaine de fixation à l'ARN, elle s'accroche à une autre protéine (Lodish *et al*, 1993 ; Suzuki *et al*, 1995).

1.25 Structures tridimensionnelles de l'ARN et de ses complexes

Les nombreuses recherches concernant les ARN ne concernent pas la nature modulaire linéaire des ARN, mais plutôt la structure tridimensionnelle. En effet, l'ARN se lie à plusieurs protéines au cours de leur vie, et se moule en des structures tridimensionnelles régulières dont la fonction est d'exprimer l'information génétique.

Les domaines de l'ARN se structurent surtout parce que les bases y sont appariées selon les mêmes règles qui stabilisent l'ADN. Quand un brin d'ADN revient sur lui-même, les régions complémentaires s'associent en une courte hélice double pour former des tiges boucles ou des boucles en épingle à cheveux. Cela s'applique également à l'ARN et lui confère les mêmes formes.

L'ARN cellulaire est en majeure partie conjuguée à des protéines pour former des complexes dont les plus gros et plus remarquables sont les ribosomes. Plusieurs recherches rapportent que des petites structures ribonucléoprotéiques appelées petites ribonucléoprotéines nucléaires "snurps" sont indispensables à la maturation des molécules d'ARN messager et à d'autres fonctions cellulaires. Enfin on connaît aussi

de petites ribonucléoprotéines cytoplasmiques comme la particule de reconnaissance du signal, qui est en jeu lors la sécrétion protéique (Lodish *et al*, 1993).

1.26 Le complexe ribonucléoprotéique hétérogène (hnRNPs)

Les précurseurs d'ARNm s'associent à de grandes quantités de protéines nucléaires comportant des domaines de fixation à l'ARN bien conservés. Ils sont associés à toute une série de protéines nucléaires, en proportion comparable à celles des histones dans les cellules qui se divisent. En traitant des extraits de cellules humaines non irradiées par des anticorps monoclonaux dressés contre plusieurs de ces protéines de hnRNPs, on identifia une série hétérogène des plus importantes d'entre elles, désignée A1-U, dont la taille varie de 34 à 120 kDa. En étudiant les ARNm qui codent, on s'aperçut que certaines sont parentes (par exemple A2 et B1) car issues d'un épissage alternatif d'exons d'une même unité de transcription.

Les protéines de particules hnRNPs se fixent très avidement à l'ARN, chacune pourtant avec sa spécificité propre. On les a groupées selon qu'elles s'attachent au poly-A, au poly-C, au poly-G ou au poly-U, ce qui suggère que chaque groupe reconnaît les différentes régions d'un ARN précurseur d'après leur composition en bases et leur séquence. Ainsi, les protéines A1, C et D contenues dans les particules hnRNPs s'attachent de préférence aux séquences riches en pyrimidines de l'extrémité 3' des introns (Lodish *et al*, 1993) Comme c'est le cas des facteurs de transcription, la plupart de ces protéines possèdent une structure modulaire, à savoir un ou plusieurs domaines de fixation à l'ARN et au moins un autre domaine sensé entrer en interaction avec d'autres protéines. En introduisant des délétions dans ces protéines de hnRNPs et en mesurant l'affinité des protéines mutées envers l'ARN, on y a détecté divers domaines de fixation à l'ARN.

À cause de leur diversité, ces protéines de hnRNPs auraient encore d'autres rôles à jouer ; certaines, par exemple, entreraient en interaction avec diverses régions d'un

pré-ARNm pour former les structures reconnues par certains facteurs de maturation des ARN. Certaines protéines de hnRNPs franchissent la barrière nucléocytoplasmique dans les deux sens, alors que d'autres ne quittent pas le noyau. Cette découverte suggère que certaines de ces protéines restent ancrées à l'ARNm parfaitement mature quand il passe du noyau au cytoplasme, pour céder ensuite leur place à des protéines cytoplasmiques, avant de regagner le noyau. Les hnRNPs sont probablement les protéines liant l'ADN simple brin les mieux caractérisées qui participent à la régulation de la transcription (Boyle 2001). Par exemple, la protéine hnRNP D et ses homologues peuvent activer ou réprimer la transcription d'une variété de promoteurs. Une autre hnRNP qui contient des cs-RBD, hnRNP A1, réprime la transcription du gène de la thymidine kinase (Lau *et al*, 2000). D'autres protéines végétales qui contiennent des cs-RBD, comme FMV-3b d'*Arabidopsis*, CEBP-1 de l'œillet et ACBF de tabac, sont capables de lier spécifiquement des éléments de régulation et représentent par le fait même des régulateurs potentiels de la transcription (Didier et Klee, 1992 ; Maxson et Woodson, 1996 ; Seguin *et al*, 1997).

Certains facteurs importants en trans s'assemblent donc même avant l'export du transcrit. Par exemple, il est connu que la protéine hnRNPA2 (Hoek *et al*, 1998), impliquée dans le transport du transcrit de la MBP, est aussi importante pour l'export nucléaire du transcrit. Il est possible que certaines protéines nucléaires puissent participer à la structure adéquate des signaux de transport, permettant une reconnaissance optimale des transcrits par l'appareil de transport cytoplasmique.

1.27 Structures des motifs des protéines ayant la fonction de lier l'ARN

Par la méthode RMN, on a pu montrer que le motif RRM de la protéine C des hnRNP consiste en 4 brins β et 2 hélices α . L'ARN s'associe à la face de la protéine formée par les 4 brins β et ses extrémités aminées et carboxylées. Ces protéines sont une des nombreuses familles qui possèdent des membres qui sont régulés lors de

basses températures. Ces protéines possèdent un domaine conservé de liaison aux acides nucléiques et peuvent se lier aussi bien à l'ADN qu'à l'ARN. Cette famille de protéines est impliquée dans plusieurs rôles dans le contrôle de la post-transcription dans la cellule tels l'épissage, le stockage, le déplacement, la stabilisation de l'ARN (Fedoroff, 2002 ; Siomi et Dreyfuss, 1997). Chez *Arabidopsis thaliana*, on dénombre 196 protéines contenant un domaine RRM (Lorkovic *et al*, 2002) qui est une signature caractéristique des membres de cette famille. C'est une séquence de 80 acides aminés contenant 2 motifs RNP1 et RNP2 qui sont composés respectivement de 8 et 6 acides aminés (Hughes et Dunn, 1996 ; Albà, et Pagès 1998; Maruyama *et al*, 1999).

Il y a également 26 protéines qui possèdent un domaine KH. Le domaine KH a un motif d'environ 60 acides aminés dont une séquence conservée de résidus hydrophobes (Grishin, 2001). La protéine K des hnRNP comporte un domaine de fixation à l'ARN appelé motif KH (pour K homology). On l'a maintenant retrouvé dans quelques protéines fixées à l'ARN, comme celle codée par le gène du X fragile (FMR1) ; certaines mutations de ce gène sont responsables de la forme la plus fréquente de retard mental congénital (Siomi *et al*, 1994). Environ la moitié de toutes ces protéines n'ont pas encore de fonction connue et d'autres études doivent être faites pour la compréhension plus détaillée de cette famille.

Un autre motif se fixant à l'ARN existe dans des protéines qui reconnaissent l'ARN bicaténaire (dsRBD), à l'exclusion de l'ADN bicaténaire. L'introduction de certaines délétions dans la protéine U des hnRNP a permis de mettre en évidence une séquence de 26 résidus comprenant cinq triades RGG (Arg-Gly-Gly) séparées par plusieurs résidus d'acides aminés aromatiques. La séquence consensus de ce motif est appelé boîte RGG, mais on n'en a pas encore découvert la structure.

1.28 La stabilité des ARNm cytoplasmiques varie beaucoup d'un type à l'autre

La concentration d'un ARNm dépend à la fois de sa vitesse de synthèse et de sa vitesse de dégradation. À vitesse de transcription égale, un ARNm à longue demi-vie atteindra des taux stationnaires bien supérieurs de ceux atteints par un ARNm rapidement dégradé après son entrée dans le cytoplasme. La stabilité d'un ARNm fixe aussi la rapidité avec laquelle la cellule peut mettre fin à la synthèse de la protéine qu'il code. Lorsque l'ARNm est stable, sa protéine est synthétisée plus longtemps après que la transcription du gène ait cessé. La plupart des ARNm bactériens sont instables, ils disparaissent de façon exponentielle avec une demi-vie de quelques minutes. Ceci permet à la cellule de réapproprier rapidement la synthèse de ses protéines aux modifications survenues dans le milieu environnant. Par contre, la plupart des cellules des organismes pluricellulaires baignent dans un milieu de propriétés constantes et remplissent une série propre de fonctions pendant des périodes pouvant s'étendre sur plusieurs mois, sinon sur la vie entière de l'organisme (exemple des neurones) ; aussi les ARNm d'eucaryotes supérieurs ont des demi-vies de plusieurs heures.

Néanmoins, il arrive que des protéines de cellules eucaryotes ne doivent agir que pendant de courtes périodes et ne s'exprimer que par moments précis. C'est le cas des hormones appelées lymphokines, qui coordonnent les interactions entre les cellules en jeu dans la réponse immunitaire chez les mammifères (Lodish *et al*, 1993). De même, les nombreux facteurs de transcription qui régissent l'enclenchement de la phase S du cycle cellulaire ne sont synthétisés que pendant de brèves périodes. La demi-vie des ARNm qui codent ces protéines est anormalement courte. Leur expression est régie par une commutation marche - arrêt rapide de la transcription de leur gène ; comme leurs ARNm sont rapidement dégradés, ces facteurs sont synthétisés à des moments précis. Bon nombre de ces ARNm de courte vie portent une série de séquences AUUUA, souvent chevauchantes, dans la région 3' non traduite. Il suffit d'insérer ces séquences riches en AU dans la région 3' non transcrite

de gènes codant des ARNm stables, tel le gène de globine, pour rendre très instables les ARNm recombinant correspondants. Le mécanisme impliqué dans la déstabilisation des ARNm par ces séquences est inconnu (Lodish *et al*, 1993).

1.29 La dégradation de certains messagers est sujette à une régulation

Dans certains cas, la vitesse de dégradation de tel ou tel ARNm est soumise à une régulation. Ainsi quand on cultive du tissu mammaire de rate en présence de prolactine, la concentration des ARNm codant la caséine du lait est d'environ 30,000 molécules par cellule. En absence de prolactine, le taux d'ARNm de caséine n'atteint que le centième de cette valeur, soit 300 molécules par cellule. L'étude in vitro des transcrits naissants a montré qu'un traitement par la prolactine d'un tissu mammaire en culture n'accélérait que trois fois la transcription du gène de caséine. On doit donc attribuer l'énorme augmentation du taux d'ARNm de caséine à l'influence de la prolactine sur la stabilité de l'ARNm de caséine.

Cette régulation de la stabilité des ARNm dans les cellules d'eucaryotes est due à un mécanisme que l'on comprend bien dans le cas de l'ARNm codant pour le récepteur de transferrine. Chez les vertébrés, le fer absorbé circule dans le sang sous forme liée, chélaté à une protéine appelée transferrine. En absence de transferrine, la trop faible solubilité d'ions de fer ne répond pas aux besoins de la cellule. Le complexe fer- transferrine entre dans la cellule en s'attachant à un récepteur particulier de la surface cellulaire. La concentration de fer dans la cellule est ajustée avec grande précision. Cet ajustement est nécessaire car un excès de fer produirait des radicaux hydroxyles extrêmement réactionnels, exerçant divers effets toxiques sur la cellule. D'autre part, l'apport en fer doit suffire à saturer les sites des multiples enzymes qui utilisent le fer comme cofacteur. Quand la cellule a du fer en réserve, elle réduit son incorporation de complexe fer - transferrine, en accélérant la dégradation des ARNm de récepteur de transferrine (TfR), ce qui réduit le taux de récepteur de transferrine. Dès que les réserves de fer s'épuisent, l'ARNm de TfR

devient plus stable et la synthèse de récepteur de transferrine augmente (Lodish *et al*, 1993).

L'étude des mutations par délétion dans le gène TfR a révélé que cet ajustement de la stabilité des ARNm de TfR dépend de certaines séquences répétées, nommées *éléments sensibles au fer* (IRE). Cette partie se trouve dans la région 3' non traduite de l'ARNm. Chaque IRE est longue d'environ 30 bases et peut se replier en une épingle à cheveux dont la boucle comporte cinq bases particulières. Dans les ARNm de TfR, la tige de ces IRE contient des séquences riches en AU, comme le sont les séquences AUUUA non transcrites qui déstabilisent les ARNm de lymphokines. Quand la cellule est bien approvisionnée en fer, ces séquences riches en AU induiraient la dégradation de l'ARNm de TfR par un mécanisme analogue à celui qui détruit les ARNm de lymphokines. Dès que la concentration du fer cellulaire diminue, une protéine cytoplasmique appelée protéine se liant au IRE va s'attacher sur les IRE en reconnaissant la séquence de la boucle ainsi que la structure de la tige. On pense que l'occupation des IRE par cette protéine empêche la reconnaissance des séquences déstabilisantes riches en AU par les protéines sensées dégrader l'ARNm de TfR. Ces données sur l'ARNm de TfR s'appliqueraient à d'autres ARNm de stabilité modulée, dont la dégradation serait ralentie par le contact d'éventuels éléments sensibles avec des protéines singulières (Lodish *et al*, 1993).

1.30 Les différents sous-groupes de protéines liant l'ARN

Comme mentionné, les protéines liant l'ARN ont un rôle prépondérant dans la régulation de la transcription des gènes. Cette régulation implique le transport des ARN du noyau, leurs localisations, leurs maturations, leurs stabilités ainsi que leurs traductions. Les protéines liant l'ARN possèdent plusieurs sous-groupes. Ces dernières ayant chacun un rôle précis lors de la régulation de la transcription. Par contre, plusieurs d'entre eux ne sont pas encore caractérisés. Il y a plusieurs protéines ayant le domaine de liaison à l'ARN (RRM) qui possèdent des domaines différents dans la

portion N-terminale de la protéine. La liste des différents domaines est indiquée au tableau 2 en annexe. Il y a également les protéines liant l'ARN qui possèdent un domaine riche en glycine du côté N-terminale. Ce sous-groupe est important puisque plusieurs études démontrent une relation entre ces protéines et plusieurs stress abiotiques comme le froid et la chaleur (Lorkovic et Barta, 2002). En plus d'être régulés sous contrôle circadien, quelques protéines de ce sous-groupe sont aussi responsables de la maturation des ARNr (Lorkovic et Barta, 2002 ; Albà, et Pagès, 1998 ; Staiger, 2001). Du fait qu'ils sont impliqués dans les stress environnementaux, et qu'il y a des membres qui sont sous rythme circadien (Harmer *et al*, 2000), les protéines liant l'ARN possédant un domaine riche en glycine pourraient être des régulateurs de CBF. Des régulateurs qui pourront expliquer l'expression différentielle de plusieurs transcrits du groupe CBF chez le blé d'hiver et de printemps.

Le but du travail est de contribuer à une meilleure compréhension des mécanismes de régulation lors de la réponse au froid chez les plantes tolérantes. Un des buts est de savoir comment les protéines riches en glycine sont régulées, et comment elles influencent l'acclimatation au froid chez le blé. Par des études de liaisons, nous voulons expliquer pourquoi il y a une accumulation de transcrits différentiels de CBF1. S'il y a stabilisation des transcrits CBF1, nous pourrions expliquer la régulation des gènes COR qui sont un groupe de gènes nécessaires pour l'acclimatation des plantes à basses températures (Sarhan et Danyluk, 1998 ; Thomashow, 1999). Bien que plusieurs éléments de cette réponse soient connus, il manque encore plusieurs pièces au casse-tête. L'approche génétique utilisée par plusieurs groupes amène beaucoup d'informations. Par contre, plusieurs expériences biochimiques essentielles démontrant la régulation de ces gènes et leurs fonctions sont manquantes.

1.31 Problématique

Les céréales représentent une des cultures les plus importantes sur la planète. Elles sont cultivées sur plus de 50 % des terres agricoles. Les baisses de température limitent la croissance, la survie et la distribution géographique des plantes. Cette limitation occasionne d'énormes pertes économiques dans des pays nordiques comme le Canada. Bien qu'une pression de sélection se soit exercée au cours de plusieurs centaines d'années, les baisses brusques de température sont fatales même si les plantes tolérantes au gel ont développé une capacité à s'acclimater aux baisses de température. La tolérance au gel prend du temps à se mettre en marche et n'est donc pas efficace pour un changement rapide de plusieurs degrés. Par exemple, pour le blé la période nécessaire pour atteindre une tolérance optimale est de 49 jours (Danyluk, 1997).

Pour la compréhension de ce phénomène complexe qu'est l'acclimatation au froid, il faut en premier lieu déterminer quels sont les gènes impliqués dans ce processus. Plusieurs études démontrent un grand nombre de gènes impliqués illustrant ainsi la complexité multigénique de la tolérance au froid. Ces différents gènes sont impliqués dans plusieurs fonctions de la cellule et leur induction à des moments de stress nous laissent croire qu'ils ont un rôle prépondérant pour la survie de la plante lors d'une baisse de température. Les facteurs de transcription sont un exemple où l'on observe une régulation lors d'une baisse de température. Les CBFs régulent les gènes COR de façon différentielle entre le blé d'hiver et de printemps. Les transcrits codant pour les CBFs pourraient être plus stables dans des cultivars de blé d'hiver pour pouvoir augmenter le niveau d'expression des gènes COR absolument nécessaires pour l'acclimatation du blé.

Cette stabilisation, pourrait provenir d'une famille de protéines liant l'ARN. C'est pourquoi l'étude de cette famille de protéine est utile afin de voir si des transcrits importants comme ceux des facteurs de transcription induits au froid sont stabilisés,

ce qui pourraient expliquer l'expression différentielle entre les cultivars de printemps et d'hiver. Cette famille de protéines est impliquée dans plusieurs aspects du contrôle post-transcriptionnel tels l'épissage, le stockage, le déplacement et la stabilisation de l'ARN (Siomi et Dreyfuss, 1997).

Chez *Arabidopsis thaliana*, on dénombre 196 protéines contenant un domaine RRM (Lorkovic et Barta, 2002) qui est une signature caractéristique des membres de cette famille. Environ la moitié de toutes ces protéines n'ont pas encore de fonction connue. D'autres études sont nécessaires pour la compréhension de cette famille qui compte plusieurs gènes dont les fonctions sont complètement inconnues ou prédites par des études d'homologies.

Une meilleure compréhension de la famille des protéines liant l'ARN et plus particulièrement ceux riches en glycines permettrait de comprendre le rôle de cette famille lors de la régulation des gènes impliqués à des stress abiotiques. De plus, cette étude permettra de voir si les protéines liant l'ARN riches en glycines lient des transcrits du facteur de transcription (CBF1). L'étude de cette régulation permettra de mieux comprendre le phénomène de la tolérance au gel et ainsi d'expliquer l'expression différentielle des gènes modulés au froid dans les deux cultivars de blé.

2.0 Identification et caractérisation moléculaires des protéines liant l'ARN lors de l'acclimatation au froid chez le blé (*Triticum aestivum* L.)

Auteurs : Michel Hecheima, Armand Miguel et Fathey Sarhan

Département des Sciences Biologiques, Université du Québec à Montréal, C.P.8888
Succ. Centre ville, Montréal, Québec H3C 3P8, Canada

2.1 RÉSUMÉ

L'acclimatation au froid et la tolérance au gel chez les céréales est un phénomène complexe et multigénique. Le blé s'acclimata au froid mais les différents cultivars ont une tolérance au gel variable. Les cultivars de blé d'hiver ont la capacité de résister à une baisse de température allant jusqu'à -27 °C contrairement aux cultivars de printemps. Les gènes modulés par le froid (COR) et leurs régulateurs CBF (Cis repeat Binding Factor) sont des gènes à la base de cette différence de tolérance. Les génotypes de blé qui accumulent le plus les transcrits de CBF1 montrent une plus grande tolérance au froid. Les bases moléculaires expliquant les différences d'expression de CBF sont inconnus. Cela peut être dû à des modifications post-transcriptionnelles ou une plus grande stabilité des transcrits. Cette stabilité peut être due aux protéines liant l'ARN. Afin de comprendre le rôle de ces protéines dans la régulation de l'ARN cibles, nous avons isolé quatre gènes appartenant à la famille des protéines liant l'ARN et caractérisé en détail une protéine liant l'ARN riche en glycine. Le gène à l'étude, *TaGr-rbp*, code pour une protéine de 162 acides aminés et est constitué de 36 % de glycine. Ce gène a été choisi car une recherche dans la banque de données de GenbankTM a montré une forte homologie en protéines (94%) avec l'orge (Num. accession S71453) qui est impliqué lors d'une baisse de température. Des études d'expression ont montré aussi que le transcrit de *Tagr-rbp* s'accumule lors d'une baisse de température, lors d'un choc thermique et lors de stress salin. Nous avons démontré également que le transcrit est présent au même niveau chez plusieurs génotypes de tolérance variable indiquant que c'est un gène impliqué lors d'une baisse de température, mais pas corrélé avec la tolérance au gel. Le gène *TaGr-rbp* a été localisé sur le chromosome 5. Une analyse par immunobuvardage montre que l'accumulation des protéines est similaire chez les variétés de blé Manitou et Norstar jusqu'à 98 jours d'acclimatation. Par contre, aucune trace de cette protéine n'est détectée chez le blé non acclimaté. Cela signifie que cette protéine joue un rôle important lors de l'acclimatation au froid. D'autres études ont été réalisées sur des lignées isogéniques et présentent le même profil protéique que la lignée parentale. Des tests de liaison montrent une interaction entre la protéine TaGR-RBP et le transcrit de CBF1. Par contre aucun test de spécificité n'a pu être réalisé. L'analyse phylogénique a permis de classer les gènes isolés du blé dans le sous-groupe des protéines liant l'ARN riche en glycine. Des études de fonction pour surexprimer le gène *TaGr-rbp* chez les plantes transgéniques sont présentement en cours afin de déterminer la fonction.

2.2 INTRODUCTION

La régulation de l'expression génique chez les organismes vivants est régulée à plusieurs niveaux. Elle peut s'effectuer au niveau de la transcription ou de la post-transcription et elle joue un rôle crucial dans la croissance et le développement. La régulation post-transcriptionnelle regroupe l'épissage de pré-ARNm, l'ajout de la structure de coiffe, la polyadénylation, le transport de l'ARN, la stabilité de l'ARN et la traduction. (Higgins, 1991 ; Dreyfuss *et al*, 1993 ; Simpson et Filipowicz, 1996). La régulation est souvent faite directement par des protéines liant l'ARN (RBP) ou indirectement lorsque ces protéines module la fonction d'autres facteurs de régulation. La large variété des cibles potentielles de l'ARN implique un grand nombre de protéines liant l'ARN qui possèdent différentes spécificités (Lorkovic et Barta, 2002). Les RBP les plus abondantes chez l'humain sont les ribonucléoprotéines nucléaires hétérogènes (hnRNPs). Elles sont définies selon leurs associations avec les transcrits naissants de l'ARN polymérase II. Le clonage de quelques gènes codant pour les hnRNPs a permis la découverte de plusieurs motifs impliqués dans la liaison à l'ARN.

Plusieurs motifs caractérisent cette famille. Le motif englobant le plus grand nombre est nommé motif de reconnaissance de l'ARN (RRM). Ce motif contient 2 courtes séquences consensus RNP1 et RNP2 qui sont un octamère et un hexamère respectivement. Ce motif est incorporé dans une séquence moyennement conservée d'approximativement 80 acides aminés. Le domaine contenant ce motif se retrouve du côté N-terminal et une grande variété de motifs auxiliaires dans la partie C-terminale (Tableau 1). Ces motifs peuvent être caractérisés entre autres par des signatures telles que riche en glycine, riche en arginine, des séquences répétitives SR ou RD. (Kenan *et al*, 1991 ; Fukami-Kobayashi *et al*, 1993 ; Burd et Dreyfuss, 1994 ; Lorkovic et Barta, 2002). Le motif RRM n'est pas seulement présent dans les protéines hnRNPs, mais aussi dans une grande variété de protéines liant l'ARN

impliquées dans toutes les régulations post-transcriptionnelles possibles. Un autre motif appelé KH (pour K homology) est considéré comme le deuxième domaine le plus fréquemment retrouvé dans la famille des protéines liant l'ARN. Ce domaine est composé d'environ 60 acides aminés et il est caractérisé par un patron composé de résidus hydrophobes conservés. La signature de cette séquence est VIGXXGXXI et elle se trouve en plein milieu du domaine. Les domaines RRM et KH sont d'anciennes structures contenues dans les protéines puisqu'ils se retrouvent dans plusieurs organismes allant de la bactérie jusqu'à l'humain. D'autres motifs contenant des doigts de zinc (zinc fingers), des jointures de zinc (Zinc knuckles) des boîtes RGG sont aussi des possibles attaches à l'ARN. Par contre, aucun de ces motifs ne peut être utilisé pour prédire une fonction possible chez la plante (Lorkovic et Barta, 2002).

Les protéines qui contiennent le motif RRM du côté N-terminal et une région riche en glycine dans la partie C-terminale (protéine liant l'ARN riche en glycine GR-RBP) sont omniprésentes chez les plantes. Les GR-RBP impliquées lors de stress abiotiques ont été étudiées lors de plusieurs tests d'expression. Depuis que le premier gène codant une GR-RBP a été identifié chez le maïs (Gomez *et al*, 1988), plusieurs autres ADNc codant pour des homologues ont été trouvés chez plusieurs espèces. (Mundy et Chua, 1988 ; Hirose *et al*, 1993 ; van Nocker et Vierstra, 1993 ; Carpenter *et al*, 1994 ; Molina *et al*, 1997 ; Moriguchi *et al*, 1997 Horvath et Olson, 1998). La plupart des études faites jusqu'à présent ont porté sur l'accumulation des transcrits et non sur l'accumulation de protéines. Il n'y a pas d'évidence que ces protéines augmentent lors de stress. Bien que la fonction de cette famille de protéine ne soit pas encore définie, plusieurs études proposent que quelques membres de cette famille aient un rôle lors de la réponse de stress biotique et abiotique. Il y a accumulation de leurs transcrits lors d'exposition au froid, à la déshydratation et à des métaux lourds et lors de blessures et infection virale. (Sachetto-Martins *et al*, 2000). Des études de localisation proposent des hypothèses de fonction de GR-RBP. Il est possible qu'elles

soient impliquées dans le métabolisme de l'ARN, plus spécifiquement dans l'épissage des ARN ribosomiques puisqu'elles sont localisées dans le nucléoplasme. D'autres études d'hybridation "in situ" ont révélé une concentration plus élevée d'ARNm dans les tissus méristématiques de jeunes plantes de maïs. L'ensemble de ces résultats suggère que les GR-RBP sont impliquées dans l'épissage de l'ARNr et le développement tissulaire des plantes. Elles peuvent aussi réguler ou affecter le développement lors de stress environnementaux.

Des études ont caractérisé un homologue de GR-RBP chez les cyanobactéries. Cette protéine ayant un motif RRM fut la première à être identifiée chez cet organisme. Ce type de protéine liant l'ARN représente une ancienne structure et semble être apparu avant la divergence des eucaryotes et procaryotes. Le rôle de la régulation des gènes par les GR-RBPs chez les plantes n'est pas encore connu puisque aucune étude de fonction *in vivo* ou *in vitro* n'a été reportée. Selon plusieurs recherches, ces protéines sont régulées selon un rythme circadien chez les plantes en plus d'être impliquées lors de plusieurs stress abiotiques dont la baisse de température. Des études réalisées par Thomashow montre qu'un facteur de transcription appelé CBF1 serait un régulateur global (master switch) pour l'activation de gènes COR (cold regulated gene). Les protéines CBF1 et d'autres membres de la famille CBF sont régulées selon un rythme circadien et impliquées lors de l'acclimatation au froid en régulant l'induction des gènes COR chez les plantes. Des études d'expression ont été réalisées et montrent que CBF1 et d'autres facteurs de transcription de la même famille sont exprimés différemment entre différents génotypes de blé de tolérance variable. Les génotypes de blé qui accumulent le plus ces transcrits montrent aussi une plus grande tolérance au gel. Peu d'études expliquent cette différence d'expression de CBF chez les céréales. Cela peut être dû à des modifications post-transcriptionnelles mais aussi à une plus grande stabilité des transcrits. Nous proposons l'hypothèse qu'une stabilisation des transcrits

de CBF1 par des protéines liant l'ARN riche en glycine pourrait être la raison qui explique ce profil d'expression différentielle chez les céréales tolérantes. Nous sommes intéressés à déterminer le rôle de la protéine riche en glycine liant l'ARN chez le blé au cours de l'acclimatation au froid.

Dans cette article, nous avons également regroupé tous les membres de la famille de protéines liant l'ARN contenant un domaine RRM (Tableau I) chez *Arabidopsis* et chez le riz afin d'avoir toutes les séquences de cette famille et de faire une étude phylogénique. Nous avons d'ailleurs isolé et séquencé 4 gènes ayant le motif RRM et des domaines auxiliaires riche en glycine. Parmi ces gènes, nous avons choisis une protéine liant l'ARN riche en glycine induite lors d'une baisse de température pour déterminer une fonction possible, puisque à ce jour aucune fonction précise est désigné.

2.3 MATÉRIELS ET MÉTHODES

2.3.1 Matériel végétal et conditions de croissance

Les cultivars de blé d'hiver Norstar et de blé de printemps Manitou ont été choisis pour les extractions d'ARN total et des protéines totales. D'autres cultivars (Clair, Quantum, Fredrick, Glenlea) ont aussi été utilisés afin de comparer génotypiquement l'expression des gènes à l'étude et de confirmer les résultats obtenus avec les blés d'hiver et de printemps utilisés lors de l'expérience. Les graines sont germées dans un mélange de terre noire et de terre promix selon une proportion (50/50) sous une longue photopériode (16 heures lumière et 8 heures noirceur) à un taux d'humidité de $75 \pm 5\%$ et une température de 21 ± 1 °C. Après une semaine de germination, les plantes ont été transférées dans une autre chambre ayant exactement les mêmes conditions énumérées plus haut à l'exception de la température qui est ajustée à 4°C pour amorcer le processus d'acclimatation. Les périodes d'exposition à basse température varient de 0 (témoin) à 98 jours. Ces étapes ont été faites pour les 2 cultivars sélectionnés soit Manitou (blé de printemps) et Norstar (blé d'hiver). Les études d'expression ont été exécutées 7, 35, 49, 77 et 98 jours de traitement au froid.

2.3.2 Extraction d'ARN et de protéines totales

L'ARN de plantes acclimatées et témoins a été extrait avec le réactif Tri Reagent en suivant les protocoles du fabricant (SIGMA). L'ARN total a été utilisé lors des études d'expression par northern blot. Les protéines solubles sont obtenues en broyant 1 g du matériel végétal, après le temps d'exposition au froid choisit, dans du tampon Tris-HCl 50 mM pH 7.5 et du DTT 10 mM. Une centrifugation à $12\,000 \times g$ a été exécutée pour éliminer les débris cellulaires. Le surnageant contenant les protéines solubles est ajouté dans du tampon de chargement (Laemli 1X final) et les échantillons sont congelées à -20°C jusqu'au moment de l'électrophorèse (SDS-PAGE).

2.3.3 Analyses bioinformatiques

Dans l'article de Lorkovic et Barta (2002), des numéros d'accessions sont donnés pour tous les membres représentant la famille des protéines liant l'ARN chez *Arabidopsis thaliana*. Tous ces numéros d'accessions permettront de choisir le sous-groupe qui nous intéresse selon la fonction donnée. En comparant par BLAST les séquences d'*Arabidopsis* présentes dans NCBI avec celle de la banque (FGAS) de EST (Expressed sequence Tags) de blé, nous avons déterminé des gènes homologues chez le blé. Dans cette banque de EST, 4 membres qui font partie des protéines liant l'ARN ont été choisis et séquencés pour obtenir leurs ADNc complets. Les clones choisis sont transformés par électroporation dans les bactéries *E. coli DH5* \square électrocompétentes. Quelques colonies ont été prélevées du plat de pétri ampicilline résistant et ensemencées dans du milieu de culture. Une purification plasmidique a été nécessaire pour l'obtention d'une quantité suffisante d'ADN. Une mémoire bactérienne avec 20 % de glycérol final a aussi été réalisée pour une conservation à -80°C. Après avoir classé toutes les séquences en nucléotides et en acides aminés des différents membres des protéines liant l'ARN d'*Arabidopsis thaliana*, il y a eu obtention de séquences homologues du riz provenant de la banque de données NCBI. Cela s'est fait par un programme automatisant les blasts réalisé au laboratoire.

2.3.4 Northern Blot

Les échantillons d'ARN total de plantes (7,5µg) ont été séparés sur gel dénaturant d'agarose à 80 volts pendant 4 heures dans du tampon MOPS 1X et transférés sur une membrane de nylon par capillarité pendant la nuit. L'ARN est fixé sur la membrane par radiation ultra-violette (U.V). Une sonde est obtenue en dénaturant les brins du gène à l'étude et en synthétisant des brins radioactifs ^{32}P à l'aide de l'ADN polymérase (Klenow). Les lavages des membranes ont été faits avec une concentration décroissante de SSC 5X 0,1% SDS. Les brins radioactifs s'hybrideront par complémentarité à l'ARNm et permettront de quantifier les transcrits du gène en question en exposant les membranes radioactives sur film Kodak Bio-max.

2.3.5 Expression de la protéine recombinante

Une fois qu'un EST ou un gène obtenu par criblage de banque d'ADNc de blé acclimaté démontre une régulation au froid par Northern blot, une protéine recombinante est produite pour des études de liaisons à l'ARN. Des oligos ont été conçus par le programme Primer3 afin de pouvoir cloner le fragment complet (partie codante) dans un vecteur d'entrée GATEWAY pENTR. Les sites de restriction déterminés par 2 enzymes différentes (BamH I et Xho I) doivent couper qu'une seule fois dans le site de polyclonage du vecteur d'entrée et aussi à l'extrémité 5' et 3' de l'ADNc. Cela permettra de cloner le fragment selon l'orientation désiré dans le plasmide d'entrée. L'amplification de la partie codante du gène *TaGr-rbp* ainsi que les sites de restriction artificiel produits par les amorces sont obtenus par un PCR selon une température d'hybridation donnée par le programme Primer 3. Le fragment obtenu est purifié sur colonne Qiagen. Une digestion et ligation avec le vecteur d'entrée Pentr 3C de Gateway permettront d'obtenir le clonage. Des bactéries *E. coli* électrocompétentes résistantes à l'antibiotique kanamycine ont été transformées avec le vecteur recombinant. Les colonies ayant incorporés l'insert de façon directionnelle pourront croître en pétries sur un milieu contenant de la kanamycine. Les plasmides possédant l'insert ont été recombinaés avec un plasmide de destination PDest 15 pour obtenir une fusion GST- N-terminal sous le contrôle d'un promoteur inductible par l'arabinose 0.1 %. Après induction avec l'arabinose, des bactéries ont été prélevées à chaque heure, jusqu'à 3 heures suivant l'induction. Les bactéries sont centrifugées à 10 000 x g pendant 10 min et le culot est resuspendu dans du tampon de solubilisation pour soniquer 10 fois pendant une période de 30 sec. Une centrifugation de 10 000 x g pendant 10 minutes est faite. Du tampon de charge (Laemli 1X final) est ajouté au surnageant et au culot et les protéines sont analysées sur un gel SDS-PAGE 12 % à 150 V pendant 1 heure. La fusion avec GST a permis la purification sur colonne d'affinité. L'analyse des fractions obtenues est réalisée sur un gel SDS-PAGE.

2.3.6 Purification de la protéine recombinante

Des billes d'agarose (1ml) liant GST (Novagen) sont déposées sur une petite colonne de 1 cm de diamètre. La résine est lavée avec 5 volumes de tampon de liaison avant de déposer les protéines totales. La colonne est lavée avec 10 volumes de tampon de liaison pour éliminer les protéines non liées à la résine suivi par l'élution avec 3 volumes du tampon d'élution 1X GST. Les fractions de 1 ml ont été recueillies dans des eppendorfs et analysées sur gel SDS-PAGE afin de vérifier la qualité de la purification. Les protéines ont été transférées sur une membrane PVDF et la bande correspondante à la protéine recombinante a été digérée par diverses peptidases et analysée par spectre de masse. Cela est nécessaire afin de s'assurer que le produit purifié est bien celui codé par l'ADNc.

2.3.7 Production d'anticorps

La protéine recombinante (150 µg) est injectée dans un lapin à l'aide de l'adjuvant Titer Max gold (Sigma). Le mélange se fait dans une proportion de 1:1 avec la protéine recombinante. Après 28 jours un rappel est fait avec 100 µg de protéine. Sept jours plus tard, le sang est récupéré et mis à 4°C pendant une nuit. Le sérum qui contient les anticorps est récupéré par centrifugation à 12 000 x g pendant 10 minutes. Un immunobuvardage est réalisé pour voir la sensibilité et la spécificité de l'anticorps. Trois rappels sont nécessaires afin de permettre au lapin de fournir des anticorps de bonne qualité et d'assurer la stabilité de l'anticorps à long terme.

2.3.8 Western Blot

Les protéines sont transférées sur une membrane PVDF et incubées dans une solution PBS tween-20 1X contenant du lait en poudre 4 %. Après 2 heures d'incubation, l'anticorps primaire anti-TaGR-RBP ou anti-GST est utilisé (Novagen). La dilution choisie est 1 : 50 000 et 1 : 20 000 respectivement. L'anticorps secondaire couplé à la peroxydase (Novagen) est utilisé pour la révélation. La dilution optimale choisie est de 1 :30 000.

2.3.9 Test de résistance au froid des bactéries

Une pré-culture bactérienne à partir de la mémoire glycérol 20 % contenant le plasmide d'expression de la protéine recombinante a été préparée jusqu'à une densité optique à 600 nm de 0.600. Il y a eu ensuite induction de la protéine recombinante avec 0.1 % d'arabinose. Une culture non induite de la même construction sert de contrôle pour l'expérience. Après 2 heures d'induction de la protéine, les bactéries sont transférées dans des bouteilles stériles pour une centrifugation de 10 minutes à 10 000 x g. Les bactéries induites et non induites sont ensuite resuspendues dans 50 ml de LB contenant 0.1 mg/ml d'ampicilline. Cette étape est cruciale pour éliminer l'arabinose des cultures induites. Une autre centrifugation est exécutée et les bactéries resuspendues dans les mêmes conditions que mentionnées précédemment. Les densités optiques des différentes cultures sont notées à l'aide d'un spectrophotomètre à 600 nm. Les milieux sont tous dilués dans une solution LB pur contenant 0.1 mg/ml d'ampicilline pour obtenir une densité 0.150. La dilution utilisée peut varier entre 1/6 et 1/10 de la densité bactérienne initiale. Une fois les dilutions bactériennes obtenues, des cultures induites et non induites sont incubées à 4°C pendant 18 heures avant le transfert à 37°C. Cela permet de quantifier la reprise de croissance bactérienne et voir si la protéine recombinante a un effet sur la croissance des bactéries au froid. D'autres dilutions de densité optique de 0.150 de cultures bactériennes non induites et induites sont incubées en parallèle à 37°C sans exposition au froid. Cela sert de contrôle et indique la croissance normale des bactéries sans stress de température. La même expérience est faite avec une autre protéine de type de MADS-Box. Ce contrôle permet de comparer l'effet de différentes protéines sur les bactéries exposées au froid.

2.3.10 Transcription in vitro

L'ADNc correspondant à la copie complète de l'ARNm du gène à l'étude est cloné dans un vecteur (pCMV-SPORT6) qui possède un promoteur spécifique d'une ARN polymérase placée en amont de l'ADNc. Après linéarisation du vecteur par une enzyme ne digérant qu'à un site spécifique, la polymérase SP6 (Promega) sera

utilisée pour transcrire "in vitro" un ARNm spécifique à une protéine qui s'appelle *cis* binding factor 1 (CBF1). L'objectif principal de cette manipulation est d'obtenir une quantité suffisante d'ARNm radioactif marqué au phosphore ^{32}P à partir d'un ADNc cloné. L'ARNm synthétisé est extrait et purifié à l'aide d'une extraction phénol:chloroforme:isoamyl. (25 :24 :1). La phase aqueuse est ensuite prélevée pour une purification additionnelle sur une colonne microspin G25. Cette étape est nécessaire afin d'éliminer les nucléotides non incorporés. L'ARNm radioactif après cette étape est suffisamment purifié pour procéder au test de liaison par l'entremise de la technique de northwestern blot et du retard sur gel. La purification est vérifiée sur un gel d'agarose 1 %.

2.3.11 Northwestern blot

Des protéines totales extraites à partir du blé Norstar ainsi que la protéine recombinante sont séparées sur gel SDS-PAGE 10 %. Les protéines sont renaturées en mélangeant doucement le gel SDS-PAGE pendant 30 minutes dans le tampon TNED (10 mM Tris-HCl pH 7.5, 50 mM NaCl, 0.1 mM EDTA et 1mM dithiothreitol) et transférées ensuite sur une membrane nitrocellulose pendant 1 heure à 100 V. Après le transfert, la membrane est incubée pendant la nuit dans du lait en poudre 4 % préparé dans un tampon 50 mM Tris-HCl pH 7.5, 50 mM NaCl, 0.1 mM EDTA et 1mM dithiothreitol. Après un lavage de 10 minutes avec du tampon TNED, des transcrits d'ARN radiomarqués (^{32}P) sont ajoutés. Le nombre de CPM d'ARN radioactifs recommandé est entre $1,0 \times 10^5$ et $1,5 \times 10^6$ par membrane. L'incubation dure entre 2 et 3 heures à la température de la pièce avec une légère agitation. Après l'hybridation, la membrane est lavée 3 fois pendant 10 min avec le tampon TNED. La membrane radioactive est exposée sur film Kodak Bio-Max pendant une période variant selon l'intensité du signal.

2.3.12 Test de liaison de l'ARN par retard sur gel

En synthétisant l'ARNm de CBF1, il y a incubation de ce transcrit radioactif par des protéines totales de plantes acclimatées, de plantes non acclimatées, et de la protéine recombinante. Le retard sur gel causé par une protéine qui lie l'ARNm sera évalué. Les réactifs de la réaction de reconstitution ont été incubés (20 µl), pendant 45 minutes à 30°C. Le milieu de reconstitution contient le tampon de reconstitution, de l'ARN radiomarké (10 000 cpm/ µl), RNasin 20 U, ATP, ARNt, et une quantité de protéine variant selon la réaction. Tout le milieu est exempt de Rnase en lavant tout le matériel avec une solution de 100 mM NaOH 0,1 % SDS après un rinçage à l'eau nanopure. L'ARN radiomarké permettra de suivre la migration de l'ARN sur un film. L'héparine a été ajoutée après l'incubation de 45 minutes afin d'éviter les liaisons non spécifiques avec l'ARN. La concentration d'héparine peut varier de 5 mg/ml à 500 µg/ml. Après une incubation de 10 minutes, 3 µl du milieu réactionnel est mise sur gel d'agarose 0,7 %. La migration se fait à 80 volts pendant une période de 4 heures. Le gel est séché pendant 2 heures pour ensuite être exposé sur film. Les détails des conditions expérimentales sont décrits dans l'article de Thomson *et al*, 1999 et de Dragon *et al*, 2000.

2.4 RÉSULTATS

2.4.1 Analyse moléculaire et structurale des séquences d'ADNc des protéines liant l'ARN

À partir de l'article de Lorkovic et Barta (2002) qui regroupe tous les gènes codant pour des protéines liant l'ARN chez *Arabidopsis*, il y a sélection de quelques numéros d'accèsion codant pour des gènes connus pour leur implication lors de stress chez les plantes. Ces gènes choisis codant pour des protéines liant l'ARN riches en glycine et une hnRNP puisque nous étions intéressés à étudier ces 2 sous-groupes de la famille dues à leur implication lors de divers stress abiotiques. Par une étude d'homologie (BLAST) de la banque de NCBI et la banque d'EST de FGAS, nous avons sélectionné des homologues chez le blé. Les numéros d'accèsions représentant ces gènes ont été sélectionnés puisque chacun des EST de la banque FGAS se trouve dans un stock bactérie-glycérol. Ces EST qui sont clonés dans le vecteur pCMV•SPORT6 sont inoculés à partir des séquences de mémoire bactérienne et une préparation plasmidique est réalisée pour l'obtention de ces EST. Ces séquences d'EST proviennent des banques d'ADNc de blé acclimaté durant 7 jours. Dans le cadre de ce projet, plusieurs EST ont été séquencés complètement et nous avons obtenu 4 EST représentant des ADNc complets codant pour des protéines liant l'ARN. Ces gènes nommés *TaGr-rbp*, (*Triticum. aestivum* L. Glycine rich RNA binding protein) *TaGr-rbp2* (*Triticum. aestivum* L. Glycine rich RNA binding protein 2), *WhGrp1* (wheat glycine rich ARN binding protein) et *Ta-hnrnp* (*T. aestivum* hétérogène Ribonucléoprotéine nucléaire) ont été alignés et corrigés à l'aide du programme bioinformatique "cap assemble".

Le premier gène, *TaGr-rbp* est représenté à la Figure 1A. Le cadre de lecture de ce gène est de 489 pb et code pour une protéine de 162 acides aminés. Son codon d'initiation débute au nucléotide 76 et le codon d'arrêt au 561^{ème}. La masse moléculaire de la protéine TaGR-RBP prédite est de 15,9 kDa avec un pH

isoélectrique théorique de 6.32. La protéine a une composition de 58 acides aminés glycine, ce qui constitue 36 % des acides aminés totaux de la protéine. Une recherche dans la banque de données de GenbankTM montre une forte homologie en protéines (94%) avec l'orge (Num. accession S71453). D'autres homologues sont retrouvés chez la plante à fleur *Rumex obtusifolius* avec un pourcentage d'homologie de 73% en acides aminés (Num. accession CAD29693.1) et le riz *Oryza sativa* avec un pourcentage d'homologie de 75% en acides aminés. (Num. accession AAT85299.1). Le domaine RNP1 composé de 8 résidus et RNP2 de 6 résidus sont soulignés en bleu sur la Figure 1A. Le motif RRM (RNA recognition motif) est du côté N-terminal de la protéine. Ce dernier est illustré sur la Figure 2A.

Les analyses des domaines trouvés à la Figure 2 ont été réalisées par Inter proScan. Ce logiciel a permis d'identifier la localisation du motif RRM sur chaque gène ainsi que le domaine RNP1 et RNP2 responsable de la liaison à l'ARN.

Le gène *TaGr-rbp2* est représenté à la Figure 1B. Le cadre de lecture de ce gène est de 498 nucléotides et code pour une protéine de 166 acides aminés. Son codon d'initiation débute au 132^{ème} nucléotide et le codon d'arrêt au 631^{ème}. La masse moléculaire de la protéine TaGR-RBP2 prédite est 16,4 kDa avec un pH isoélectrique théorique de 7.96. La protéine a une composition de 42 acides aminés Glycine, ce qui constitue 25 % des acides aminés total de la protéine. Une recherche dans la banque de données de GenbankTM montre une homologie en protéines (69%) avec le riz (Num. accession NP_914833.1). D'autres homologues sont retrouvés comme chez le pois *Pisum sativum* avec un pourcentage d'homologie de 56 % en acides aminés (Num. accession AAB71417.1) et le tabac *Nicotiana glauca* avec un pourcentage d'homologie de 53% en acides aminés. (Num. accession BAA11089). Le domaine RNP1 composé de 8 résidus et RNP2 de 6 résidus sont soulignés en bleu sur la Figure 1B. Le motif RRM (RNA recognition motif) est situé au milieu de la protéine et cette dernière est illustrée à la Figure 2B.

Le gène WhGrp1 est représenté à la Figure 1C. Ce gène code pour une protéine liant l'acide nucléique simple brin déjà isolé chez *Triticum aestivum*. Le cadre de lecture de ce gène est d'une longueur de 507 nucléotides et la séquence prédite en protéine de 167 acides aminés. La masse moléculaire est de 16,3 kDa et avec un pH isoélectrique théorique de 5.18. Son codon d'initiation débute au 121^{ème} nucléotide et le codon d'arrêt au 625^{ème}. La protéine a une composition de 68 acides aminés glycine, ce qui constitue 41 % des acides aminés totaux de la protéine. Une recherche dans la banque de données de GenbankTM montre une homologie en protéines de 87% avec l'orge (*Hordeum vulgare*) (Num. accession CAA88558.1). D'autres homologues sont retrouvés comme chez le riz *Oryza sativa* avec un pourcentage d'homologie de 78 % en acides aminés (Num. accession AAT85299.1) et le sorgho *Sorghum bicolor* avec un pourcentage d'homologie de 73% en acides aminés (Num. accession CAA40862.1). Le domaine RNP1 composé de 8 résidus et RNP2 de 6 résidus sont soulignés en bleu sur la Figure 1C. Le motif RRM (RNA recognition motif) est situé au milieu de la protéine et cette dernière est illustrée à la Figure 2C.

Le dernier gène, Ta-hnRNP est représenté à la Figure 1D. Le cadre de lecture de ce gène est d'une longueur de 1365 nucléotides et code pour une protéine de 454 acides aminés. Son codon d'initiation débute au nucléotide 73 et le codon d'arrêt au 1435^{ème}. La masse moléculaire de la protéine Ta-hnRNP prédite est de 45.3 kDa avec un pH isoélectrique théorique de 6.18. La protéine a une composition de 109 acides aminés Glycine, ce qui constitue 22 % des acides aminés totaux de la protéine. Une recherche dans la banque de données de GenbankTM montre une homologie (82%) en protéines avec un membre de riz *Oryza sativa* non étudié (Num. accession AAP54226.1). D'autres homologues sont retrouvés aussi comme chez *Arabidopsis thaliana* avec un pourcentage d'homologie de 54% en acides aminés (Num. accession NP_193166.2) et une autre protéine ribonucléique chez *Arabidopsis* (Num. accession CAB10209.1). Le domaine RNP1 composé de 8 résidus et RNP2 de 6 résidus sont soulignés en bleu sur la Figure 1D. Il y a 2 motifs RRM (RNA recognition motif) qui

sont localisés du côté N-terminal de la protéine et cette dernière est illustrée sur la Figure 2D. Les 2 domaines RRM sont soulignés en jaune et la portion riche en glycine est en gris.

Différents organismes possèdent des domaines RRM. La Figure 3 illustre de façon schématique les principaux domaines des protéines liant l'ARN chez des organismes différents. Les bactéries (cyanobactéries), les plantes supérieures (blé) ainsi que les mammifères (souris), montrent une structure semblable malgré leurs grandes divergences évolutives. Les domaines RRM sont plus souvent localisés du côté N-terminal et la partie riche en glycine du côté C-terminal. La longueur de la portion riche en glycine varie d'une espèce à l'autre lorsqu'on compare les domaines riches en glycine de différentes protéines. De plus, aucune corrélation ne peut être faite sur la longueur et l'espèce. La longueur de ce domaine auxiliaire chez les plantes ou les mammifères varie énormément et ne peut constituer une variable pouvant aider à la détermination d'une fonction précise. Le nombre de copies du domaine RRM peut aussi varier et ce même chez la même espèce. Les protéines liant la queue Poly (A) de l'ARN chez les plantes possèdent quatre copies du domaine RRM alors que d'autres sous-groupes ne possèdent que 2 copies (ribonucléoprotéine hétérogène nucléaire). Il peut y avoir qu'une seule copie comme dans le cas des protéines riches en glycine ou des protéines impliquées dans l'épissage alternatif. La longueur des différents motifs de ces protéines est détaillée à l'échelle.

2.4.2 Domaines auxiliaires retrouvés chez la famille des protéines liant l'ARN

Le tableau I (en annexe) représente quelques domaines auxiliaires qui se trouvent du côté C-terminal de différentes protéines liant l'ARN. Ces domaines auxiliaires sont toujours présents dans tous les sous-groupes de cette famille et sont directement liés à des rôles importants quant à la capacité de la protéine à lier l'ARN, à sa spécificité, et à ses différentes fonctions possibles. On voit quelques exemples de domaines se trouvant sur les protéines liant l'ARN qui lui confèrent certains rôles

métaboliques. À ce jour, plusieurs domaines ne sont pas encore connus avec certitude, mais certains peuvent être utilisés pour hypothétiser certaines fonctions des protéines qui possèdent ces domaines.

2.4.3 Recherche des numéros d'accension d'*Arabidopsis thaliana* et de *Oryza sativa*

Lors de cette étude, nous avons recueilli tous les numéros d'accension d'*Arabidopsis thaliana* afin de pouvoir avoir tous les membres de protéines liant l'ARN et de les classer en sous groupes, de pouvoir faire des études phylogéniques sur ces membres et de pouvoir trouver leurs homologues chez un autre organisme végétal qui est lui aussi complètement séquencé, le riz (*Oryza sativa*). Ces séquences obtenues d'*Arabidopsis thaliana* et de *Oryza sativa* pourront également nous aider à identifier une région conservée entre deux espèces et concevoir ainsi, après alignements des séquences, des oligos spécifiques afin d'obtenir l'homologue chez le blé. À l'aide d'oligos spécifiques, une amplification d'un gène d'intérêt chez le blé peut être obtenu à partir d'une banque d'ADNc.

Le tableau II (en annexe) représente tous les numéros d'accessions des protéines d'*Arabidopsis thaliana* ayant un motif RRM (1^{ère} colonne). Toutes les protéines ayant un motif RRM du génome d'*Arabidopsis* ont été répertoriées. Chaque membre a été comparé par TBLASTN pour trouver son homologue chez le riz. La deuxième colonne de ce tableau montre 2 membres chez le riz ayant la plus forte homologie avec celui d'*Arabidopsis thaliana*. Leurs rôles hypothétiques ou la définition du gène donné par GenbankTM est également mentionné sur cette même colonne. La dernière colonne nous indique la valeur d'homologie trouvée par NCBI lorsque ces blasts ont été réalisés. Une valeur proche de 0 pour le « E value » montre une homologie significative alors que les valeurs entre E^{-1} et E^{-26} ne sont pas très significatives et une analyse visuelle des séquences trouvées s'impose.

2.4.4 Expression du gène *TaGr-rbp* lors de l'acclimatation au froid et de stress abiotiques

Le gène *TaGr-rbp* a été choisi puisque son homologue chez l'orge était exprimé au froid, et que des études d'expression faites au laboratoire avec le blé montraient son association avec les baisses de température. De plus, ce gène est localisé sur le chromosome 5 du blé qui est reconnu comme porteur de plusieurs gènes régulateurs de l'acclimatation chez le blé (Figure 7). Pour déterminer le rôle du gène *TaGr-rbp* lors de l'acclimatation au froid, une étude d'expression par l'analyse northern blot a été nécessaire. L'étude a permis de déterminer que le transcrit s'accumule de façon plus prononcée chez le blé et les espèces céréalières acclimatées pendant 36 jours comparativement à ceux non acclimatés au froid (Figure 4). Le génotype Norstar (blé d'hiver) qui est le cultivar le plus tolérant démontre une nette accumulation du transcrit par rapport au même condition de croissance des plantes non acclimatés. Seule l'avoine qui est très sensible au froid n'accumule pas le transcrit *Tagr-rbp*. L'accumulation par contre est semblable chez les différents génotypes, ce qui laisse croire que ce transcrit est induit au froid mais n'est pas directement relié à la tolérance au gel.

Une étude d'expression en réponse aux stress environnementaux a aussi été faite. Les transcrits de *Tagr-rbp* sont accumulés lors de plusieurs stress abiotiques. L'expression est légèrement présente en condition de chaleur intense (40°C), de stress salin, et lors de baisse de températures. Il faut noter que l'expression de *TaGr-rbp* est inhibé lors de sécheresse comparativement au contrôle puisque lors d'une déshydratation (45 % de la masse initiale en eau), il n'y a moins d'accumulation que le témoin. La différence entre le témoin FTA1 et WS CTRL sur la Figure 5 est que FTA1 a été récolté lorsque le blé était en terre et mis à -80°C. Le WS CTRL a été déraciné du pot pour un lavage des racines et ensuite congeler à -80°C. Le témoin WS CTRL pour la sécheresse est nécessaire puisque les plantes soumis à ce stress sont traitées différemment du témoin FTA1. Il n'y a aucune accumulation lors de

traitement à l'acide abscisique. Tous ces tests d'expression ont été réalisés avec le génotype Fredrick qui est un blé d'hiver. La dernière étude d'expression montre que l'accumulation de *Tagr-rbp* est croissante du début de l'acclimatation jusqu'à la fin pour les 2 géotypes à l'étude (Figure 6).

2.4.5 Production et purification de la protéine TaGR-RBP

La protéine recombinante de TaGR-RBP a été produite par les bactéries *E. coli* BL21 1A. La Figure 8 montre son induction pendant une période de 3 h. La Figure 8 B démontre par immunobuvardage que la protéine de fusion avec un Tag GST est bien présente et que les bactéries produisent la protéine. L'intensité du signal démontre qu'une heure d'induction est suffisante aux bactéries pour produire une quantité suffisante de la protéine de fusion GST- TaGR-RBP. La solubilité de la protéine est égale dans les 2 parties (surnageant et culot) puisqu'après sonication des bactéries, on retrouve un signal identique dans les deux fractions.

La protéine recombinante GST- TaGR-RBP possède une masse moléculaire de 41.9 kDa puisque la protéine native est de 15.9 kDa et que l'attache GST est de 26 kDa. La purification s'est faite sur une colonne contenant une résine qui lie les molécules ayant une attache GST. L'immunobuvardage réalisé sur les différentes fractions démontrent une excellente purification puisqu'aucune trace de la protéine n'est retrouvée dans les échantillons de lavage et qu'une infime partie se retrouve non attachée dans les échantillons obtenues lors du chargement des protéines sur colonne (Figure 9).

2.4.6 Immunobuvardage de la protéine TaGR-RBP par l'anticorps polyclonal

L'anticorps anti- TaGR-RBP obtenu par immunisation d'un lapin permet de détecter la protéine native dans différents échantillons de blé acclimaté. La Figure 10 A et B montrent l'abondance de la protéine TaGR-RBP chez la lignée parentale de Manitou et Norstar. Des protéines totales de cultivar de blé Manitou et Norstar

acclimatés jusqu'à 98 jours ont été extraites pour cette étude. Par une étude d'immunobuvardage, on voit que l'accumulation de la protéine TaGR-RBP chez les 2 génotypes est presque identique et que leurs accumulations respectives augmentent avec le temps d'acclimatation. L'immunobuvardage montre l'abondance relative des protéines par rapport au témoin (0 jours d'acclimatation) où on n'observe aucune accumulation. Une autre étude par immunobuvardage a été réalisée sur des protéines totales solubles de lignées isogéniques et le résultat montre une accumulation identique avec les lignées parentales (Figure 10 C et D). Cette expérience démontre que la localisation du gène *TaGr-rbp* ne se trouve pas sur le locus *Vrn-A1* (Figure 7).

2.4.7 Test de liaison

Des tests de liaison par northwestern blot nous montrent une interaction du transcrit de CBF1 (l'ARNm de CBF1 radioactifs) avec la protéine recombinante TaGR-RBP (Figure 11). Il n'y a aucun signal observé aux points servant de contrôle négatif comme le BSA et celui représentant les protéines totales bactériennes non induites. Il n'y a aucun signal observé lors des autres conditions d'expérimentations à l'exception de la protéine purifiée. Par contre, un très faible signal est observé dans le puit Ca total et Ca 40-60% de sulfate d'ammonium. Ces résultats montrent qu'il y a une possible interaction, mais cette expérience ne permet pas de connaître la spécificité de la liaison du transcrit envers sa protéine (Figure 11).

2.4.8 Étude phylogénique

Toutes les séquences de la famille des protéines liant l'ARN d'*Arabidopsis* contenant le motif RRM ont été récupérées et alignées par le programme clustal X afin de procéder à l'arbre phylogénique. Les 4 séquences de blé des gènes *TaGr-rbp*, *TaGr-rbp2*, *WhGrp1*, *Ta-Hnrhp* ont été inclus dans l'alignement, et on voit après la génération de l'arbre que les protéines de blé se retrouvent dans une sous-famille où les numéros d'accension d'*Arabidopsis* correspondent au même sous-groupe (Figure 12).

2.5 DISCUSSION

Les ADNc de *Tagr-rbp* et des autres gènes isolés lors de cette étude ont été obtenus dans les banques d'ADNc de blés acclimatés au froid. Quatre gènes ont été isolés lors de cette étude, et le gène *TaGr-rbp* a été caractérisé plus en détail. Il contient 2 domaines distincts, le premier domaine est le motif RRM (recognition RNA motif) qui est du côté N-terminal, et le second domaine riche en glycine dans la partie C-terminale de la protéine. Le gène *TaGr-rbp* chez le blé a été choisi puisque son homologue chez l'orge était exprimé au froid. Nos études d'expression chez le blé montraient son implication lors de l'acclimatation au froid. Une recherche dans la banque de donnée de NCBI montre des homologues chez plusieurs espèces dont *Arabidopsis thaliana*, le riz, le maïs et l'orge. Jusqu'à présent aucune fonction précise n'a été avancée pour ce gène. Le gène a été localisé sur le chromosome 5 du génome du blé. Ce chromosome est reconnu comme porteur de plusieurs facteurs de transcription impliqués lors de la régulation de la tolérance au gel chez le blé. De plus, plusieurs QTLs associés avec des caractères importants pour la tolérance au gel et à d'autres stress abiotiques ont été localisés sur ce chromosome.

Plusieurs organismes possèdent des protéines contenant des domaines de liaison à l'ARN. Parmi ces derniers, le domaine RRM est l'une des principales signatures. Ces domaines sont présents chez des organismes différents comme les bactéries (cyanobactéries), les plantes, la souris, le xenope et l'humain. Plusieurs études montrent que les gènes exprimant des protéines riche en glycine liant l'ARN comme la protéine CIRP de souris et la protéine GR-RBP de cyanobactéries sont induits au froid. De plus, ces protéines sont induites à d'autres types de stress chez les plantes tels que les blessures et les métaux lourds. Dans le cas du gène *TaGr-rbp* isolé dans cette étude, en plus d'être induit au froid, son transcrit est exprimé lors de traitement de choc thermique et lors d'un stress salin. Cela montre que le sous-groupe dans lequel il appartient, est impliqué aussi dans une multitude de stress abiotique chez les plantes.

La structure de ces protéines est très semblable chez les organismes vivants (Figure 3). Les domaines RRM sont localisés du côté N-terminal et la partie riche en glycine du côté C-terminal. La longueur de la portion riche en glycine varie d'une espèce à l'autre pour la protéine TaGR-RBP, et aucune corrélation ne peut être faite sur la longueur et l'espèce. La longueur de ce domaine auxiliaire chez les plantes ou les mammifères ne varie pas énormément. Par contre, cette caractéristique ne constitue pas un paramètre pouvant aider à la détermination d'une fonction précise. Toutefois, ils ont tous en commun le fait qu'ils sont impliqués à une fonction précise lors d'une baisse de température, mais qui n'est pas encore déterminée. Ceci supporte la notion que les protéines riches en glycine liant l'ARN représentent une classe impliquée dans la réponse moléculaire aux basses températures.

Pour déterminer le rôle du gène *TaGr-rbp* lors de l'acclimatation au froid, une étude d'expression par analyse northern a été nécessaire. L'étude a permis de déterminer que le transcrit s'accumule de façon plus prononcée chez le blé et les espèces céréalières acclimatées comparativement à ceux non exposés au froid (Figure 4). L'accumulation par contre est presque constante chez les différents génotypes, ce qui laisse croire que ce transcrit est impliqué lors de l'acclimatation au froid mais n'est pas directement relié à la tolérance au gel. Ceci s'explique du fait que l'accumulation du transcrit n'est pas directement proportionnelle au niveau de tolérance des différents génotypes étudiés. Des études de longues cinétiques au froid ainsi qu'à différents stress ont été réalisées, et les résultats montrent que l'expression du gène *TaGr-rbp* n'est pas spécifique au froid seulement, et qu'il pourrait donc être impliqué lors d'autres types de stress abiotiques. Nous avons aussi noté qu'un traitement à l'ABA n'induit pas le gène, ce qui est un étonnant puisque les homologues de ce gène sont influencés par l'ajout exogène de cette hormone.

Nos travaux ont démontré grâce à l'anticorps anti-TaGR-RBP que cette protéine s'accumule chez Manitou et Norstar. L'accumulation est identique chez les 2

génotypes et aucune trace de cette protéine ne se trouve chez le blé non acclimaté. Ceci démontre que cette protéine est impliquée seulement lors d'une baisse de température et qu'en condition normale de température la protéine TaGR-RBP n'a aucune fonction puisqu'elle n'est pas du tout traduite. De plus on observe une accumulation croissante de la protéine au cours de l'acclimatation. Cela indique que l'implication de cette protéine lors de l'acclimatation à froid est plus importante lorsque le processus d'acclimation est avancé. Étant donné que l'accumulation de la protéine est comparable chez ces 2 cultivars, nous proposons que ce gène est impliqué au froid mais qui n'est pas corrélé à la tolérance au gel. Ce résultat correspond aux tests d'expression réalisés avec différents génotypes qui démontreraient une accumulation des transcrits au froid sans être directement relié à la tolérance au gel des différents génotypes étudiés (Figure 4). Ceci peut s'expliquer du fait que des homologues de ce gène sont présents chez plusieurs organismes qui ont un degré de tolérance au froid variable. Cela est le cas par exemple des souris, humains et des bactéries et des différents génotypes utilisés.

Nous avons déterminé également l'accumulation de la protéine TaGR-RBP chez les lignées isogéniques. Ce sont des lignées obtenues par croisement génétique et le suivi du croisement est en fonction du QTL de la vernalisation (Limin et Fowler, 2002). Le génotype Manitou ayant le comportement génétique de Norstar puisque le QTL de vernalization (*vrn*) y ait incorporé. L'accumulation protéique des lignées isogéniques Manitou et Norstar présentent le même profil d'expression protéique que la lignée parentale de Manitou et Norstar. Cela s'explique du fait que la protéine TaGR-RBP a été localisée dans une autre région que le QTL de vernalization. Le profil d'expression protéique devrait donc être identique dû au fait que la protéine n'est pas localisée sur le QTL *vrn* incorporé par croisement génétique. Ce résultat va dans le même sens que celui obtenu lors de la localisation du gène sur le chromosome 5 faite précédemment (Figure 7).

Nous avons testé l'effet de la protéine TaGR-RBP sur la croissance bactérienne lors d'une incubation à 4°C. Des études précédentes réalisées par Kim *et al*, 2005, ont montré que la protéine AF42 d'*Arabidopsis* permettait une plus grande croissance d'*E.coli* après une exposition au froid. Cette étude par Kim *et al*, démontrait un effet protecteur contre une baisse de température sur les bactéries. Etant donné que la protéine AF42 possède une structure très semblable à celle de TaGR-RBP, nous avons voulu vérifier si cet effet se retrouvait sur la protéine TaGR-RBP. Notre étude n'a pas démontré un effet protecteur sur les bactéries lors d'une baisse de température. Le témoin (contenant plasmide vide) et les bactéries induites produisant la protéine TaGR-RBP semblent croître au même rythme et aucun effet sur la croissance n'a été observé lorsque la protéine était dans le milieu. Cela veut dire qu'il n'y a aucune protection au froid par la protéine TaGR-RBP sur les bactéries.

Plusieurs études démontrent l'implication des protéines riches en glycine liant l'ARN lors de plusieurs stress abiotiques. Il y a accumulation de leurs transcrits lors d'une baisse de température, déshydratation, blessure, infection virale ainsi que la toxicité des métaux lourds (Sachetto-Martins *et al*, 2000). Des études de localisation démontrent que plusieurs GR-RBP sont situés dans le nucléoplasme. L'ensemble de ces résultats suggèrent qu'ils soient impliqués dans le métabolisme de l'ARN, et plus spécifiquement lors de la maturation des ARN ribosomiques. Des études d'hybridation "in situ" ont révélé une concentration plus élevée d'ARNm dans différents tissus méristématiques de jeunes plantes de maïs. L'ensemble de ces résultats tend à suggérer que les GR-RBP sont impliqués lors de la maturation du précurseur 45S des ARNr et le développement tissulaire des plantes en période de stress. Ils peuvent aussi réguler ou affecter ces fonctions lors de stress environnementaux (Naqvi *et al*, 1998 ; Alba et Pages, 1998 ; Lorkovic et Barta, 2002). À la lumière de ces résultats et de ces hypothèses proposées dans différentes études, nous proposons que la protéine TaGR-RBP pourrait jouer une fonction

similaire à celle des autres organismes puisque leurs structures inter-espèces est très similaire. De plus, plusieurs organismes possèdent des protéines riches en glycine liant l'ARN et elles sont également très conservées dans la même espèce (Lorkovic et Barta, 2002). Cela peut indiquer qu'elles auraient une fonction commune dans la même et dans les différents organismes. D'après nos résultats, nous suggérons que la protéine TaGR-RBP peut jouer un rôle semblable à ses homologues des différentes espèces et serait impliqué dans le métabolisme des ARN ribosomiques. Ces derniers sont très actifs lors d'une baisse de température afin de pouvoir traduire les protéines ayant un rôle direct ou indirect lors de la tolérance au gel. La protéine TaGR-RBP pourrait se lier aux ARN ribosomiques afin de pouvoir accélérer ou améliorer le rendement de la traduction de différentes protéines. Ces protéines qui sont nécessaires lors de la réponse au froid ou à un stress quelconque nécessitant une traduction rapide des protéines impliquées dans les mécanismes de défense. L'accélération ou l'amélioration du mécanisme de traduction pourraient donc être le résultat d'une meilleure stabilisation des transcrits ou des protéines ribosomiques. Une protéine liant l'ARN peut également se lier à d'autres protéines pour former des complexes plus stabilisateurs pour l'ARN et c'est pourquoi, une étude d'interaction protéine-protéine par l'entremise du système double hybride chez la levure serait utile pour confirmer notre hypothèse. Cette expérience qui est en cours permettra de déterminer les interacteurs possibles avec la protéine TaGR-RBP et de voir s'il y a une liaison avec des protéines ribosomiques.

Mise à part la caractérisation de la protéine TaGR-RBP, nous avons voulu vérifier si la protéine TaGR-RBP pouvait se lier au transcrit de CBF1 et le stabiliser. On sait que ce facteur de transcription est d'une importance capitale lors de l'acclimatation au froid en induisant les gènes COR. Un northwestern blot a été réalisé afin de tester cette liaison. Les résultats démontrent une interaction avec la protéine recombinante ainsi qu'avec la protéine purifiée (Figure 12). Malgré la liaison obtenue, ce test ne démontre aucune spécificité, c'est pourquoi, nous avons voulu refaire l'expérience à

l'aide d'un test de liaison par retard sur gel. Cette expérience n'as toutefois pas confirmé cette liaison. Par contre, le test de liaison par northwestern blot confirme que la protéine recombinante produite était dans un état natif puisqu'elle lie l'ARN radioactif. D'autres expériences de retard sur gel doivent être faites afin de confirmer ou infirmer la liaison du transcrit de CBF1.

Une analyse phylogénique par l'entremise du programme Paup 4.0 a été réalisée. Cette étude nous a permis de voir dans quel sous-groupe les gènes isolés du blé seront classés parmi ceux d'*Arabidopsis*. On sait que les gènes d'*Arabidopsis* ont été classés en sous-groupes (Lorkovic et Barta, 2002) selon leurs fonctions putatives. Comme prévu, les 3 gènes riches en glycine isolés lors de cette étude ont été classés également dans le même sous-groupe que celui d'*Arabidopsis*, c'est-à-dire les protéines riches en glycine. La protéine Ta hnRNP de blé se trouve dans le groupe des protéines nucléaires hétérogènes d'*Arabidopsis* selon l'arbre. Cela nous montre la proximité évolutive de ces deux protéines (voir numéro d'accension en gras Figure 13).

En plus d'avoir servis pour l'analyse phylogénique, les données séquentielles des gènes d'*Arabidopsis* nous ont permis par des blast automatisés d'obtenir ceux du riz (Tableau 2). Ces séquences pourraient par alignements nous permettre de trouver une portion conservée d'un gène d'intérêt et de pouvoir l'amplifier par des oligos spécifiques. Ces séquences permettront aussi de trouver les positions des exons des différents gènes. Ceci est important pour la conception d'oligos spécifiques pour l'étude d'expression utilisant la technologie quantitative real time PCR (qRT PCR). Des études de fonction pour surexprimer le gène TaGr-rbp chez les plantes transgéniques sont présentement en cours.

2.6 RÉFÉRENCES

- Alba, M.M. and Pages, M. 1998. Plant proteins containing the RNA recognition motif. *Trends Plant Sci.* 3, 15–21.
- Burd, C.G. and Dreyfuss, G. 1994. Conserved structures and diversity of functions of RNA-binding proteins. *Science*. 265, 615–621.
- Carpenter, C.D., Kreps, J.A. Simon, A.E. 1994. Genes encoding glycine-rich *Arabidopsis thaliana* proteins with RNA-binding motifs are influenced by cold treatment and an endogenous circadian rhythm. *Plant Physiol.* 104, 1015–1025.
- Dreyfuss, G., Matunis, M.J., Pinol-Roma, S. Burd, C.G. 1993. HnRNP proteins and the biogenesis of mRNA. *Annu. Rev. Biochem.* 62, 289–321.
- Fukami-Kobayashi, K., Tomoda, S. Go, M. 1993. Evolutionary clustering and functional similarity of RNA-binding proteins. *FEBS Lett.* 335, 289–293.
- Gomez, J., Sanchez-Martinez, D., Stiefel, V., Rigau, J., Puigdomenech, P. Pages, M. 1988. A gene induced by the plant hormone abscisic acid in response to water stress encodes a glycine-rich protein. *Nature*. 334, 262–264.
- Higgins, C.F. 1991. Stability and degradation of mRNA. *Curr. Opin. Cell Biol.* 3, 1013–1018.
- Hirose, T., Sugita, M. Sugiura, M. 1993. cDNA structure, expression and nucleic-acid binding properties of three RNA binding proteins in tobacco: occurrence of tissue alternative splicing. *Nucleic Acids Res.* 21, 3981–3987.
- Horvath, D.P. Olson, P.A. 1998. Cloning and characterization of cold-regulated glycine-rich RNA-binding protein genes from leafy spurge (*Euphorbia esula* L.) and comparison to heterologous genomic clones. *Plant Mol. Biol.* 38, 531–538.
- Kim YO, Kim JS, Kang H. 2005. Cold-inducible zinc finger-containing glycine-rich RNA-binding protein contributes to the enhancement of freezing tolerance in *Arabidopsis thaliana*. *Plant J.* 42(6), 890–900.
- Kenan, D.J., Query, C.C. Keene, J.D. 1991. RNA recognition: towards identifying determinants of specificity. *Trends Biochem. Sci.* 16, 214–220.

Limin A.E, Fowler DB. 2002. Developmental traits affecting low-temperature tolerance response in near-isogenic lines for the Vernalization locus *Vrn-A1* in wheat (*Triticum aestivum* L. em Thell). *Ann Bot (Lond)* 89(5),579-85

Lorkovic, Z.J. Barta, A. 2002. Genomic analysis: RNA recognition motif (RRM) and K homology (KH) domain RNA-binding proteins from the flowering plant *Arabidopsis thaliana*. *Nucleic Acids Res.* 30, 623–635.

Molina, A., Mena, M., Carbonero, P. Garcia-Olmedo, F. 1997. Differential expression of pathogen-responsive genes encoding two types of glycine-rich proteins in barley. *Plant Mol. Biol.* 33, 803–810.

Moriguchi, K., Sugita, M. Sugiura, M. 1997. Structure and subcellular localization of a small RNA-binding protein from tobacco. *Plant J.* 6, 825–834.

Naqvi SM, Park KS, Yi SY, Lee HW, Bok SH, Choi D. 1998. A glycine-rich RNA-binding protein gene is differentially expressed during acute hypersensitive response following Tobacco Mosaic Virus infection in tobacco. *Plant Mol Biol.* 37(3),571-576.

Sachetto-Martins, G., Franco, L.O. de Oliveira, D.E. .2000. Plant glycine-rich proteins: a family or just proteins with a common motif? *Biochim. Biophys. Acta.* 1492, 1–14.

Simpson, G.G. Filipowicz, W. 1996. Splicing of precursors to mRNA in higher plants: mechanism, regulation and sub-nuclear organization of the spliceosomal machinery. *Plant Mol. Biol.* 32, 1–41.

Van Nocker, S.Vierstra, R.D. 1993. Two cDNAs from *Arabidopsis thaliana* encode putative RNA binding proteins containing glycine-rich domains. *Plant Mol. Biol.* 21, 695–699.

2.7 LÉGENDES DES FIGURES

Figure 1 : Séquence et organisation structurale des gènes codant des protéines liant l'ARN caractérisés chez le blé.

(A) Séquences d'acides aminés et nucléotidique du gène *T. aestivum* L. Glycine rich RNA binding protein (TaGR-RBP). Le cadre de lecture (ORF) est d'une longueur de 489 nucléotides et la séquence prédite en protéine est de 162 acides aminés. La masse moléculaire est de 15,9 kDa avec un pH isoélectrique théorique de 6.32. Le domaine RRM est souligné en jaune et la portion riche en glycine est en gris. Les nucléotides en rouge sont ceux utilisés pour la conception d'oligo pour amplification PCR lors du clonage de l'ADNc. RRM). Le motif RRM contient 2 courtes séquences consensus RNP1 et RNP2 qui sont un octamère et un hexamère respectivement sont représentés en bleu. Le codon d'initiation et le codon d'arrêt sont en majuscules. Les nucléotides en vert sont ceux qui ont été choisis pour l'obtention par PCR d'une sonde spécifique.

(B) Séquence d'acides aminés et nucléotidique du gène *Triticum. aestivum* L. Glycine rich RNA binding protein 2 (TaGR-RBP2) chez le blé. Le cadre de lecture (ORF) est d'une longueur de 498 nucléotides et la séquence prédite en protéine est de 166 acides aminés. La masse moléculaire est de 16,4 kDa et avec un pH isoélectrique théorique de 7.9. Le domaine RRM est souligné en jaune et la portion riche en glycine est en gris. Le motif RRM contient 2 courtes séquences consensus RNP1 et RNP2 qui sont un octamère et un hexamère respectivement sont représentés en bleu. Le codon d'initiation et le codon d'arrêt sont en majuscules. Cette séquence représente une protéine liant l'ARN riche en glycine non caractérisée à ce jour.

(C) Séquence d'acides aminés et nucléotidique du gène wheat glycine rich ARN binding protein (whGRP1) chez le blé. Ce gène code pour une protéine liant l'acide nucléique simple brin. Le cadre de lecture (ORF) est d'une longueur de 507 nucléotides et la séquence prédite en protéine est de 167 acides aminés. La masse moléculaire est de 16,3 kDa et avec un pH isoélectrique théorique de 5.18. Le domaine RRM est souligné en jaune et la portion riche en glycine est en gris. Le motif RRM contient 2 courtes séquences consensus RNP1 et RNP2 qui sont un octamère et un hexamère respectivement sont représentés en bleu. Le codon d'initiation et le codon d'arrêt sont en majuscules.

(D) Séquence d'acides aminés et nucléotidique du gène *T. aestivum* hétérogène Ribonucléoprotéine nucléaire (Ta-hnRBP) chez le blé. Ce gène code pour une protéine appelée Ribonucléoprotéine hétérogène nucléaire (hnRNP). Le cadre de lecture (ORF) est d'une longueur de 1365 nucléotides et la séquence prédite en protéine est de 454 acides aminés. La masse moléculaire est de 45,3 kDa avec un pH isoélectrique théorique de 6.18. Les 2 domaines RRM sont soulignés en jaune et la portion riche en glycine est en gris. Le motif RRM contient 2 courtes séquences consensus RNP1 et RNP2 qui sont un octamère et un hexamère respectivement sont représentés en bleu. Le codon d'initiation et le codon d'arrêt sont en majuscules.

Figure 2 : Représentation schématique des différents motifs des gènes isolés ainsi que leurs longueurs en acides aminés. Représentation de la protéine TaGR-RBP (A), TaGR-RBP2 (B), WhGRP1 (C) et Ta-hnRNP (D) chez le blé. Le domaine de reconnaissance de l'ARN (RRM) est représenté par un rectangle rouge et le domaine riche en glycine est représenté par un rectangle gris.

Figure 3 : Représentation schématique de la structure des protéines liant l'ARN de plusieurs organismes.

Les protéines liant l'ARN contiennent un ou plusieurs motifs RRM. La longueur des protéines est illustrée à l'échelle. Le domaine RRM est présenté en noir et le domaine riche en glycine en jaune. Le nombre de RRM est indépendant de l'espèce. Le nombre de ce dernier peut également varier dans la même espèce dépendamment du sous-groupe dans laquelle appartient la protéine.

Figure 4 : Accumulation des transcrits de *Tagr-rbp* durant l'acclimatation au froid chez différents génotypes de blé et d'espèces de céréales.

Les différents génotypes suivants ont été étudiés. Avoine cv Laurent LAU (*Avena sativa* L.); Orge cv Winchester WIN (*Hordeum vulgare* L.); 2 génotypes de blé de printemps, cv Glenlea GLEN et cv Concorde CON (*Triticum aestivum* L.); 4 génotypes de blé d'hiver, cv Monopole MON, cv Absolvent ABS, cv Norstar NOR, cv Fredrick FRED (*Triticum aestivum* L.); Seigle d'hiver cv Musketeer MUS (*Secale cereale* L.) Dans cette étude, un échantillon de 7.5 µg d'ARN total ont été séparés sur un gel dénaturant de formaldéhyde et hybridé avec une sonde du gène *TaGr-rbp* radiomarqué au ³²P. Les températures entre parenthèses représentent la température à laquelle 50% du blé meurent (LT50). Les bandes d'ARN ribosomique 18S servent pour la quantification et l'équilibre des charges de l'ARN mis sur gel.

Figure 5 : Accumulation des transcrits de *Tagr-rbp* durant l'acclimatation au froid et lors de stress abiotiques.

FTA1, contrôle non acclimaté de blé d'hiver Fredrick. FEA1, plantes acclimatées à 4°C pour un jour. FHA1 et FHA3, plantes exposés à 40°C pendant 1 et 3 heures respectivement. FW, plantes exposées à une blessure et récoltées selon le temps indiqué en heure. NA300 et NA500, plantes traitées avec 300mM et 500mM NaCl pour une période de 18 heures. WS, plantes déshydratées de 45% par rapport à sa teneur relative en eau de départ. ABA, plantes traitées avec 0.1 mM pour une période de 18 heures.

Figure 6 : Accumulation des transcrits de *Tagr-rbp* durant l'acclimatation au froid chez le blé de printemps Manitou et le blé d'hiver Norstar.

Témoin, blé non acclimaté. 7, 21, 35, 49, 77 et 98 jours, période d'acclimatation à 4°C des 2 géotypes. Les bandes d'ARN ribosomique 18S servent pour la quantification et l'équilibre des charges de l'ARN mis sur gel.

Figure 7: Localisation chromosomique de *TaGr-rbp* chez le blé (*Triticum aestivum* L.) La localisation chromosomique se fait à l'aide de la lignée ditélocentrique. Le gène *TaGr-rbp* est situé sur le bras long du chromosome 5 du blé *Triticum aestivum* L. Les locus *Vm-A1* et celui du groupe CBF sont déjà caractérisés. Ils sont représentés en bleu et jaune respectivement.

Figure 8 : Production de la protéine recombinante de TaGR-RBP chez la bactérie *E. coli*

(A) SDS-PAGE montrant l'induction de la protéine TaGR-RBP fusionnée avec GST. Les puits # 1 à 8 représentent les protéines du surnageant. Les pistes 1,3,5,7 représentent les bactéries non induites représentant le temps 0, 1 heure, 2 heures et 3 heures respectivement. Les pistes 2,4,6,8 représentent les bactéries induites représentant le temps 0, 1 heure, 2 heures et 3 heures respectivement. Les puits 9 à 14 représentent les protéines du culot. Les pistes 9,11,13 représentent les bactéries non induites représentant le temps 1 heure, 2 heures et 3 heures respectivement. Les pistes 10,12,14 représentent les bactéries induites représentant le temps 1 heure, 2 heures et 3 heures respectivement. La piste 15 représente le standard moléculaire Broadrange.

(B) Analyse d'immunobuvardage des fractions protéiques obtenues suivant l'induction des bactéries *E. Coli*. Les protéines ont été transférées sur une membrane PVDF et la détection est réalisée à l'aide d'un anticorps Anti-GST 1 :10 000. La flèche noire indique la protéine recombinante fusionnée à GST produite par la bactérie.

Figure 9 : Purification de la protéine recombinante de fusion GST- TaGR-RBP produite par induction des bactéries *E. coli* BL21 1A.

(A) Gel SDS-PAGE montrant la purification de la protéine de fusion GST- TaGR-RBP. Les puits # 1 à 3 représentent les protéines lors du chargement de la colonne. Les puits # 4 à 10 représentent les échantillons de lavage alors que les puits # 11 à 15 représentent l'élution de la protéine recombinante par le glutathion 1X.

(B) Analyse d'immunobuvardage des fractions protéiques obtenues suivant la purification des protéines des bactéries *E. coli*. Les protéines ont été transférées sur une membrane PVDF et la détection est réalisée à l'aide d'un anticorps Anti-GST 1 :10 000. La flèche noire indique la protéine recombinante fusionnée à la GST.

Figure 10: Accumulation de la protéine TaGR-RBP chez le blé de printemps (Manitou) et d'hiver (Norstar) et les lignées isogéniques lors de l'acclimatation au froid

Les protéines totales du blé acclimaté de Manitou et Norstar ont été utilisées lors de cette expérience. Après un gel de protéine SDS-PAGE selon les mêmes conditions que la Figure 7, ces protéines ont été transférées sur une membrane PVDF. La détection est réalisée à l'aide d'un anticorps Anti-TaGR-RBP 1 : 50 000. Le sérum du lapin non purifié contenant les anticorps fut utilisé pour la détection. La Figure 10A représente le signal obtenu pour le génotype Manitou et la Figure 10B représente le signal obtenu pour le génotype Norstar. Les puits représentent respectivement 0, 7, 21, 35, 49, 77 et 98 jours d'acclimatation chez le génotype Manitou et Norstar. La Figure 10C représente le signal obtenu pour le génotype Norstar-Printemps Nil et la Figure 10D représente le signal obtenu pour le génotype Manitou – Hiver Nil. Les puits représentent respectivement 0, 7, 21, 35, 49, 77 et 98 jours d'acclimatation chez le génotype Norstar-Printemps Nil et Manitou – Hiver Nil. La protéine rubisco a servi pour la quantification et l'équilibre des charges des protéines déposées sur gel.

Figure 11 : Test de liaison par northwestern blot du transcrit CBF1 radioactifs sur différents échantillons de protéines.

(A) Analyse northwestern de différentes fractions protéiques de blé et de BSA. Na total = protéines totales provenant de blé non acclimaté. Na 40-60% = protéines totales provenant de blé non acclimaté précipitées au sulfate d'ammonium selon le pourcentage indiqué. TaGR-RBP non-induit = protéines bactériennes totales. TaGR-RBP induit = protéines bactériennes totales produisant la protéine recombinante. Ca total = protéines totales provenant de blé acclimaté. Ca 40-60% = protéines totales provenant de blé acclimaté précipitées au sulfate d'ammonium selon le pourcentage indiqué. Prot. *E.coli* Total = protéine bactérienne d' *E. coli* ne contenant aucun plasmide. TaGR-RBP pure (5 µg/µl), protéine purifiée sur colonne. BSA = albumine provenant du serum de bovin, servant de contrôle négatif.

(B) Gel SDS-PAGE montrant les différentes fractions protéiques de blé et de BSA utilisées lors de du test de liaison par la méthode northwestern. Ce gel a été renaturé et hybridé avec le transcrit CBF1 radioactif.

Figure 12 : Arbre phylogénique de gènes d'*Arabidopsis* et de blé réalisé par le "programme paup 4.0 Beta"

Toutes les séquences d'*Arabidopsis* ont été récupérées et alignées par le programme clustal X afin de procéder à l'arbre phylogénique. Les 4 séquences de blé *Tagr-rbp*, *Tagr-rbp2*, *Whgrp1*, *Ta-hnrpb* ont été inclus dans l'alignement afin de voir dans quel groupe l'arbre les classera parmi ceux d'*Arabidopsis*. Notons que les numéros d'accèsion utilisés pour cet arbre sont représentatifs de toute la famille des protéines liant l'ARN et que certains homologues ont été enlevés pour alléger l'arbre.

2.8 FIGURES

```

1  cttctccccactcagtgtcggttcccgcccggttcgggttcgtagggtttagtagtgg
1      M A D V E Y R C F V G G L A W
61  ggaagaggaggaagaATGgcggaagtcgagtagcgtgcttcgtcggcggcctcgccctgg
16  A T D D Q S L Q N A F S K Y G D V I D S
121 gccaccgacgaccagtcctccagaacgccttctccaagtacggcgacgtcatcgactcc
36  K I I T D R E T G R S R G F G F V T F A
181 aagatcatcactgacagggagacgggcccgttcccgcgggttcgggttcgtcaccttcgcg
56  S D E A M R Q A I E A M N G Q D L D G R
241 tcggacgaggcgatgcgccaggcgatcgaggccatgaacggccaggacctggacggccgc
76  N I T V N E A Q S R R S G G G G G G F S
301 aacatcacggtcaacgagggccagtcctccgcccgtccggcgaggcgccggcggttcagc
96  G G G G G G Y G G Q R R E G G G G G Y G
361 ggcggaggcggtggcggttacggcgccagcgccgagggcgggcgggcggttacggc
116 G G G G G Y G S G R S G G G G G Y G S R
421 ggcggcgggcggtggctacggaagtggcgcagcgggcgggcggtggctacggcagccgt
136 E G G G G Y G G G G G Y G G S R G G S
481 gagggcgggcggttacggcgggcgggcggttacggcggcagccgcggttcg
156 G G G K W R E *
541 ggcggcggaagtggagggagTGAatggtggggccctcgctggccagttatcctagtta

601 cccgatcccgtttttacctgtctacccaaatgttatcgctcgctcgctcccaggtatcct

661 agatctcgctccgtcgcgtaggttccgagatgttcaggctaccgttagatgtgacgtgc

721 taccgttcggtgttacgttccgtccccgctctgttggctatgcttctaagattgaaatg

781 aaaacctgagcaaagtttggttcgcatgtgatgcttctctcggaagtttggttccttctggt

841 tgtgccgtgaacgcttggccttgccactgggtgatattttgccgcacccactaggtgata

901 cgtatttttggtgcagattttgtagcccctgaattcgattggctcgactgggacgtgctcc

961 tcccttaagttgatgttccccagtaaaaaaaaaaaaaaaaaa

```

Figure 1A

```

1      cgcggacgcgtggcgactggagcagaggacactgacatggactgaaggagtagaaaac

61      tccccattctccattccccctcccacgccgtctcgteccagctcgccctccagaggttt
1      M A A L I K L G S L L R Q S A L
121     tagtcttcctccATGgcggccctcatcaagctcggaagccttcttaggcagagtgccttg
17      A S S A S T G S A P A L F N A A R F M C
181     gcgagcagtgcctcgaccggctcggctcctgccctgttcaacgctgcccgttcatgtgc
37      A G Q P A G G N K L F V G G L A W A T D
241     gccggccagccagctggtggcaataagcttttcgtcggcggctctggcctgggctactgac
57      D H S L K E A F Q S F G E V V E A R V I
301     gatcattcgctgaaggaggcgttccaaagctttggagaggtcggttgaggcacgtgtcatc
77      T D R E T G R S R G F G F V S F A N E A
361     actgatcgagagactggaaggtctagagggtttggttttgtagctttgcgaatgaagct
97      D A D S A A K G M D G Q E L A G R N V R
421     gatgcagatagcgtgcaaaaggcatggacggtcaggagcttgacgggaggaacgtccgt
117     V N F A N E R P S G G G G F S R G G G G
481     gtgaacttcgccaatgagagaccttcgggtggtggcggcttctcacgcggcggtggc
137     Y G G G G G F G G G G G Y G G G G R G G
541     tatggaggaggtggtggtttggcgggcggtcggtacggcggtggtggccgaggtggc
157     G G Y G G D R E A F *
601     ggtggctatggtggtgaccgtgaagctttcTAAggagagcgcagactatcaaatttaaaa

661     taatgcgcttttaggacagtgtcttgattatgtggaacctatgtcacgtctgtctgttat

721     ctgcgatagttgaactcaataggggcagagatgtggttctagtatttttgtttttggttc

781     gtttcttgatggtgcaagagtttaagtcatgctgcttcgaggcttgatgagtcctcaatg

841     gaagttcatgcattttgccgtatattgcactgtgtacttggtgagtcctctgagttaag

901     tcgtgtatggttagtctgattaaaaaaaaaaaaaa

```

Figure 1B

```

1      ccggttggttctctctccctcccggttcgggttcgggttctggttcggttcgggtccga

61      gagtagctagcctagtctagggtttccgtagagaggaggaggaggaagagagagatcc
1      M A E T E Y R C F V G G L A W A T D D N
121     ATGgcggagacggagtagcgtgcttcgtgggcggcctcgctgggccaccgacgacaac
21      N L Q Q A F S Q Y G E I L D A K I I N D
181     aacctccagcaggccttcagccagtagcggcgagatcctcgacgccaagatcatcaacgac
41      R E T G R S R G F G F V T F G S E E S M
241     cgcgagacggggaggtcccgcggttcgggttcgtcacgttcggcagcgaggagtcgatg
61      R Q A I E E M N G K E L D G R N I T V N
301     cgccaggccatcgaggagatgaacggcaaggagctcgacgggcgcaacatcacggtcaac
81      E A Q S R R S G G G G G G G G G G G G Y G G
361     gaggcccagtcgccgctcgggcggagggggcggcggcgggcgggcggtacggcggc
101     Q R G G G G G Y G G G G G Y G G G G G G
421     cagcgcggcggtggcgaggttacggcggcggcggcggctacggaggcggcggcgggcggc
121     Y G G Q G G G G Y G G Q R G G G G Y G
481     tacggcggcaggcggtggcggtacggcggcagcgcggtggcggcgggtacggc
141     G G G G Y G G G G G Y G G Q R G G G G
541     ggcggcggcgtacggcggtggcggtacggcggcagcgcggtggcggcggc
161     D S G G Q W R N *
601     gactccggcgccagtgagggaacTGAgcctcaaattctccagttatctatctatctacc

661     ttgtcgtcgtcctcgtcgtcttcttcagcaagttatc

```

Figure 1C

1 tctccccatcgccccctcccgcaaccgagcgcgagagaatccgaggagaggagcgggcg
 1 M E S D Q G K L F I G G I S W E
 61 aaggaggcggtgATGgagtcggatcagggcaagctcttcatcgcgccatctcctgggag
 17 T T E E K L Q E H F S N F G E V S Q A A
 121 acgacggaggagaagctgcaggagcacttctccaacttcggcgaggctctccaggccgcc
 37 V M R D K L T G R P R G F G F V V Y A D
 181 gtcatgcgcgacaagctcactggcgcccggggcttcggcttcgtagtctacgccgac
 57 P A A V D A A L Q E P H T L D G R T V D
 241 cccgcccgcgtcgacgccgccctccaggagccccacaccctcgacggccgcacggctgat
 77 V K R A L S R E E Q Q A T K A V N P S A
 301 gtgaagcggcgctctcgcgggaggagcagcaggctaccaaggcggtgaaccctagcgca
 97 G R N A G G G G G G G G G G G D A G G A
 361 ggaaggaacgctggaggtggtggcgggcgcgcgcgggcgggcgcgatgccgggtggtgct
 117 R T K K I F V G G L P S S L T D E E F R
 421 aggacaaagaagatttttgtggcggaactgccctccagctctgacagatgaggagttccgg
 137 Q Y F Q T F R A V T D V V V M Y D Q T T
 481 cagtacttccagaccttcgggctgtcaccgatgttggtgatgtatgaccagacaaca
 157 Q R P R G F G F I T F D S E D A V D R V
 541 cagcgtccccggggcttcggcttcattaccttgactcggaggatgcggttgaccgtgtg
 177 L H K T F H D L G G K M V E V K R A L P
 601 ctgcacaaaaccttcacgatcttgagggaagatggttagaggtgaagcgtgctctgcc
 197 R E A N P G S G G G G R S M G G G G F H
 661 cgagaggcgaatcctggtctggcgggcgggccggttccatgggaggtgggggttttcat
 217 S N N G P N S N A S S Y D G R G D A S R
 721 agtaacaatggaccaactccaatgctagcagctatgatggcagaggcgatgctagcaga
 237 Y Q Q A Q A S G G Y P G Y G A G A Y G
 781 tatgggcaggcgagcaagccagtggtggctacccaggttatggtgctggagcttatggc
 257 S A P T G Y G Y G H T N P G T T Y G N Y
 841 agtgctccaactgggtatggatatgggcacaccaatcctggaactacatatggaaattat
 277 G S A G Y G A V P A A Y A G A Y G N P S
 901 gggctctgcagggtatggagctgttcctgctgcgtatgctggggcttatggcaatccaagt
 297 A A A S G Y Q G G P P G A N R G P W G S
 961 gctgctgcttcgggttaccagggtggccctccaggcgctaataagaggaccctggggcagt
 317 Q A P S G Y G T G G Y A G N A G Y G A W
 1021 caagctccgtctggttatggcactgggggttatgctggcaatgcaggctatggtgcatgg
 337 N S S S A G G N A P T S Q A P G A A A G
 1081 aatagctcttctgctggtgggaatgcaccactagtcaggcacctggtgcagctgcaggt
 357 Y G N Q G Y G Y G G Y G G D A S Y G N H
 1141 tatggaaaccagggtatggttatggtggatatggaggagatgcatcatatggtaatcat
 377 G G Y G A Y G G R G D G A G N P A T G A
 1201 ggagggtatggtgcttatggtggccggggagatggtgctggaaatcctgctactggtgca
 397 A S G Y G A A G Y G S G N G N S G Y P N
 1261 gcctctgggtatggtgctgctggatatggaagtgggaatggaaactctggatatccaat
 417 A W T D P S Q G G G F G G A V N G S E
 1321 gcgtggactgatccttcacaaggcggggatttggtggtgcagtcgaatggagctctgag
 437 G Q S N Y G S G Y G G M Q P R V A Q *
 1381 ggccaatcaaattatggcagtggttatggtggtatgcagcctagggttgctcagTAGaag
 1441 aggctgttataatgttcatggtgtgaaccagtgcatattgggcagcatatgtattattgcg

Figure 1D



Figure 2

Protéine liant l'ARN de cyanobactérie (bactérie)



Protéine liant l'ARN de chloroplaste (plante)



Protéine liant l'ARN riche en glycine (plante)



Protéine liant l'ARN riche en glycine (mammifère)



Ribonucléoprotéine hétérogène nucléaire (plante)



Protéine de plante pour l'épissage alternatif (plante)



Protéine de Plante liant la queue Poly(A) (plante)



100 acides aminés

Figure 3

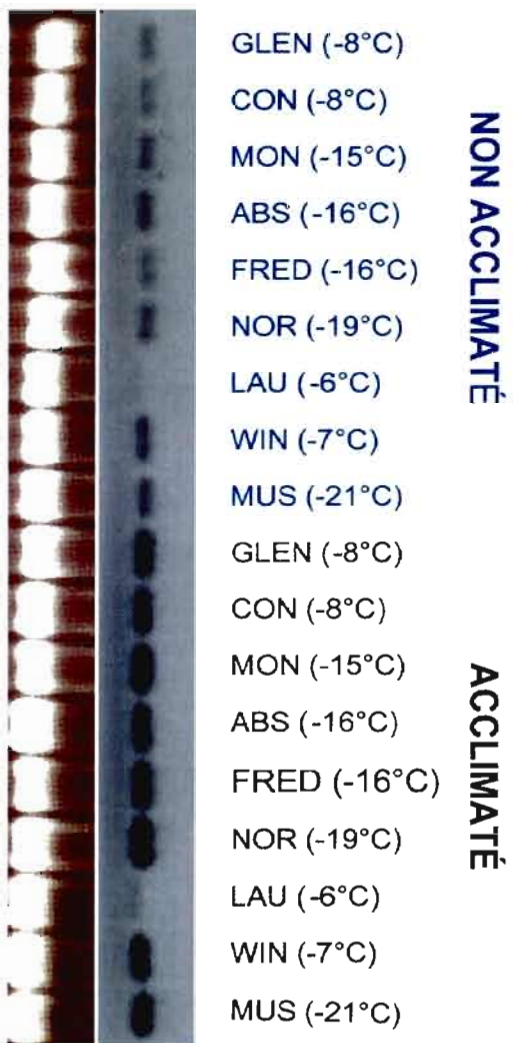


Figure 4

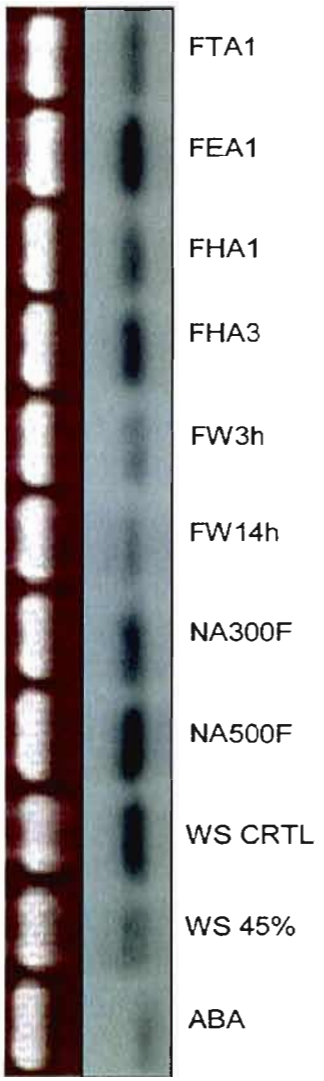
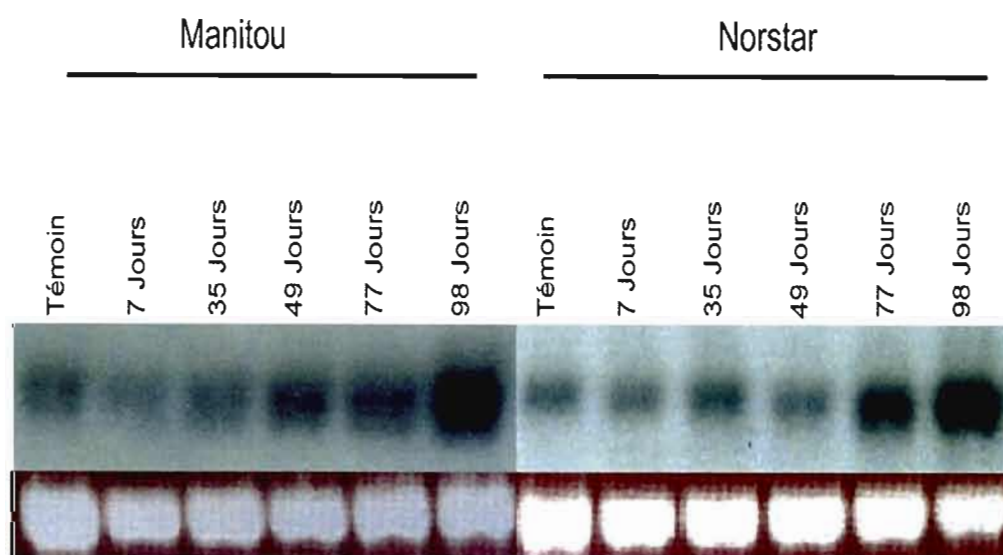


Figure 5

**Figure 6**

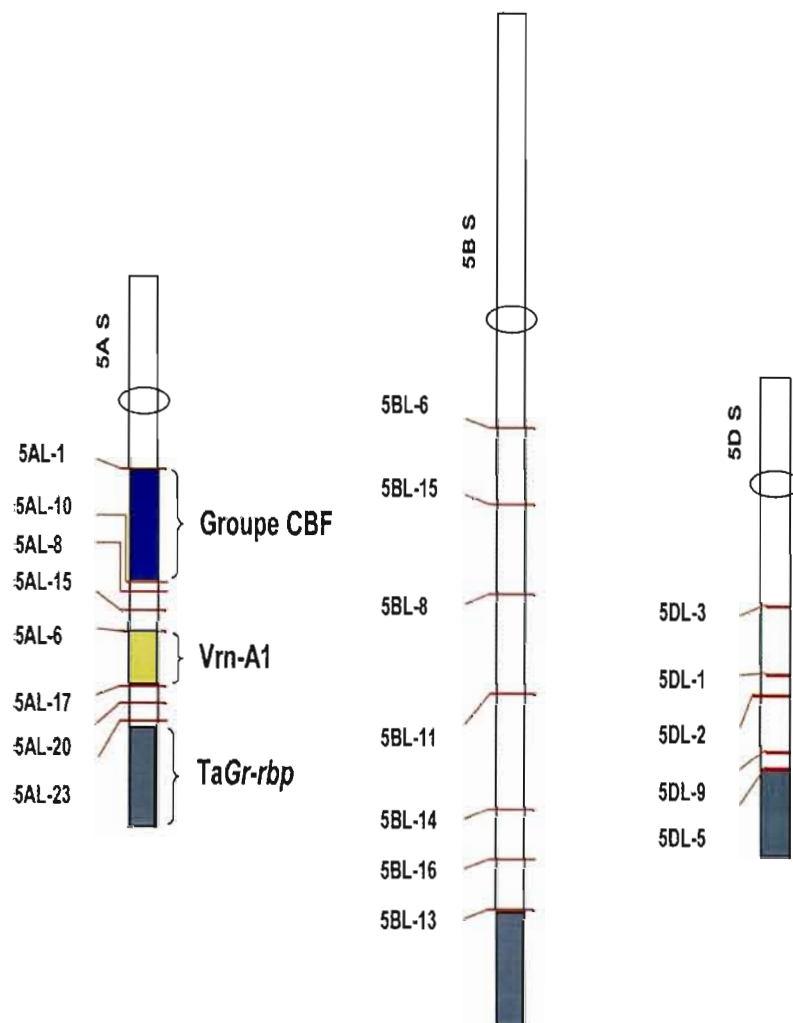


Figure 7

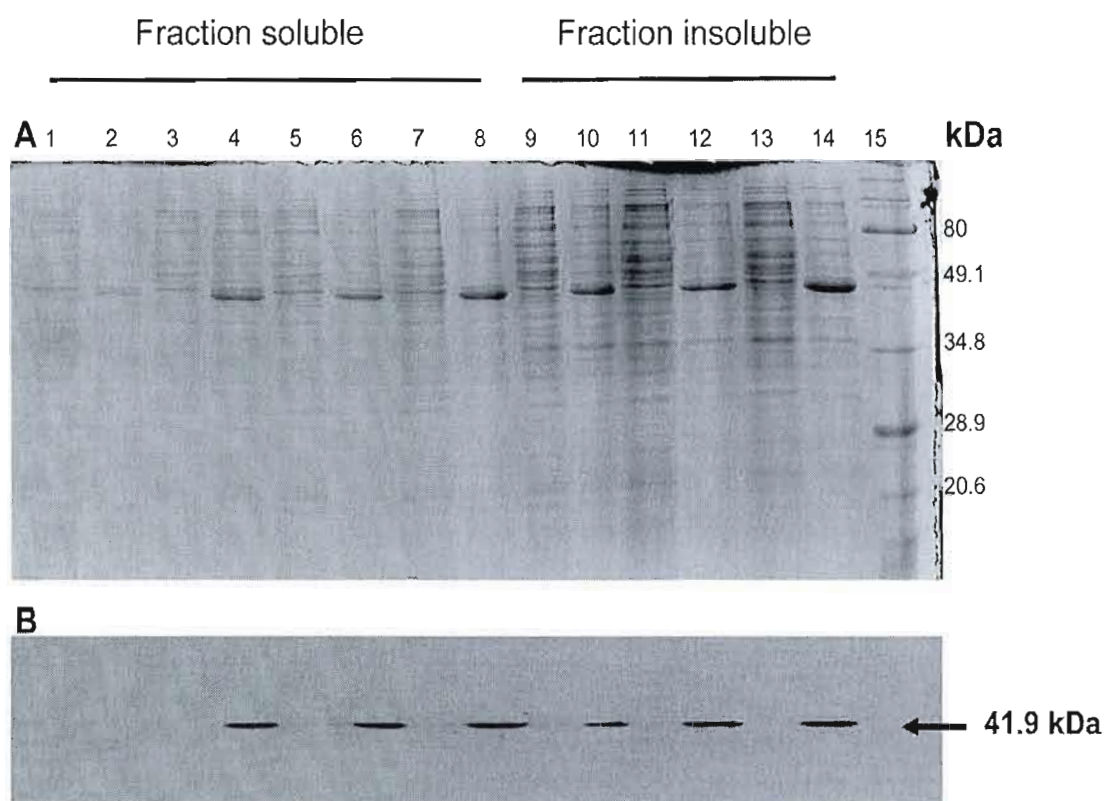


Figure 8

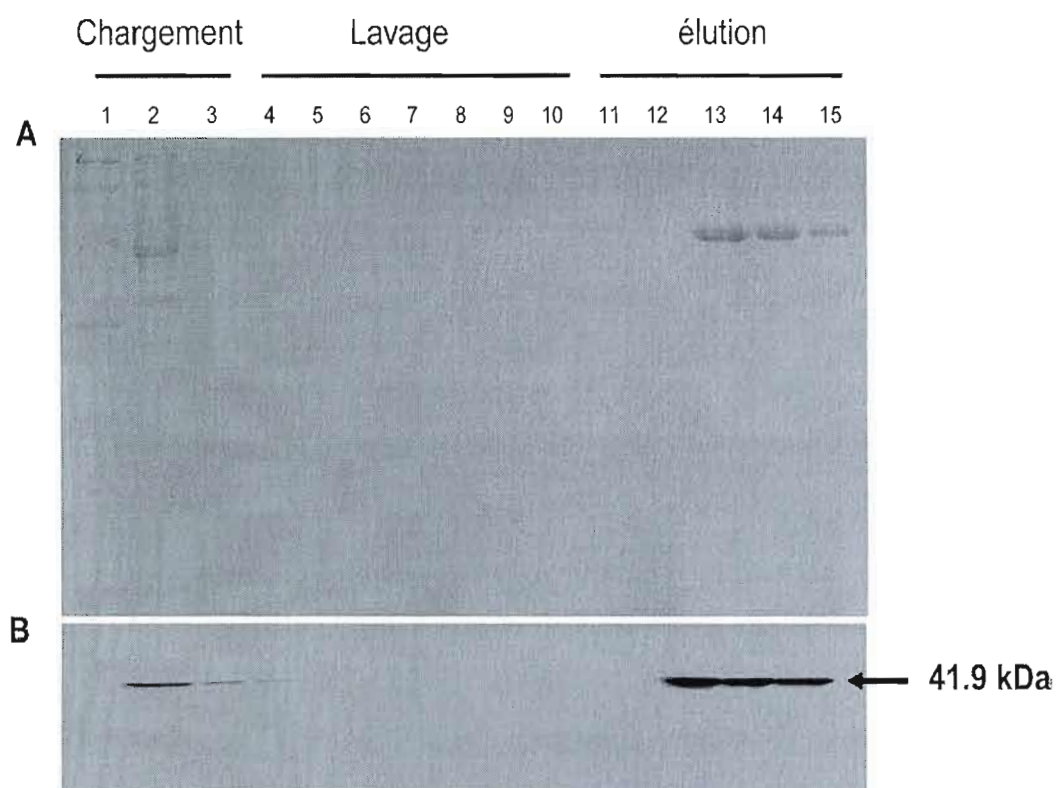
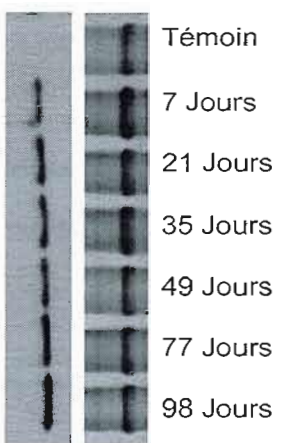


Figure 9

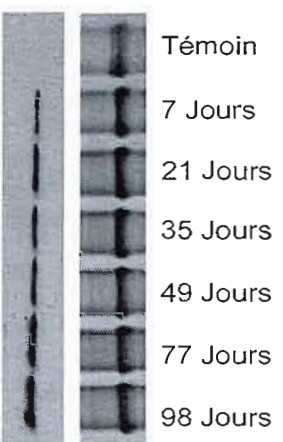
Lignée Parentale

A

Manitou

**B**

Norstar



Lignée isogénique

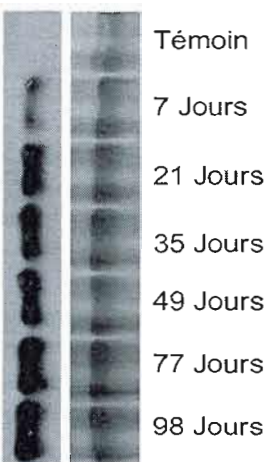
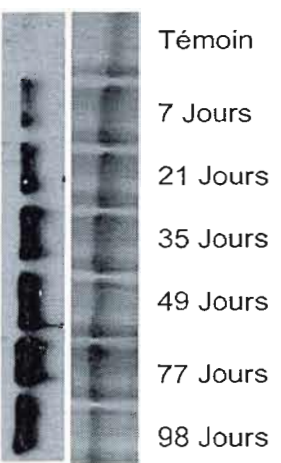
C Norstar - Printemps Nil**D** Manitou – Hiver Nil

Figure 10

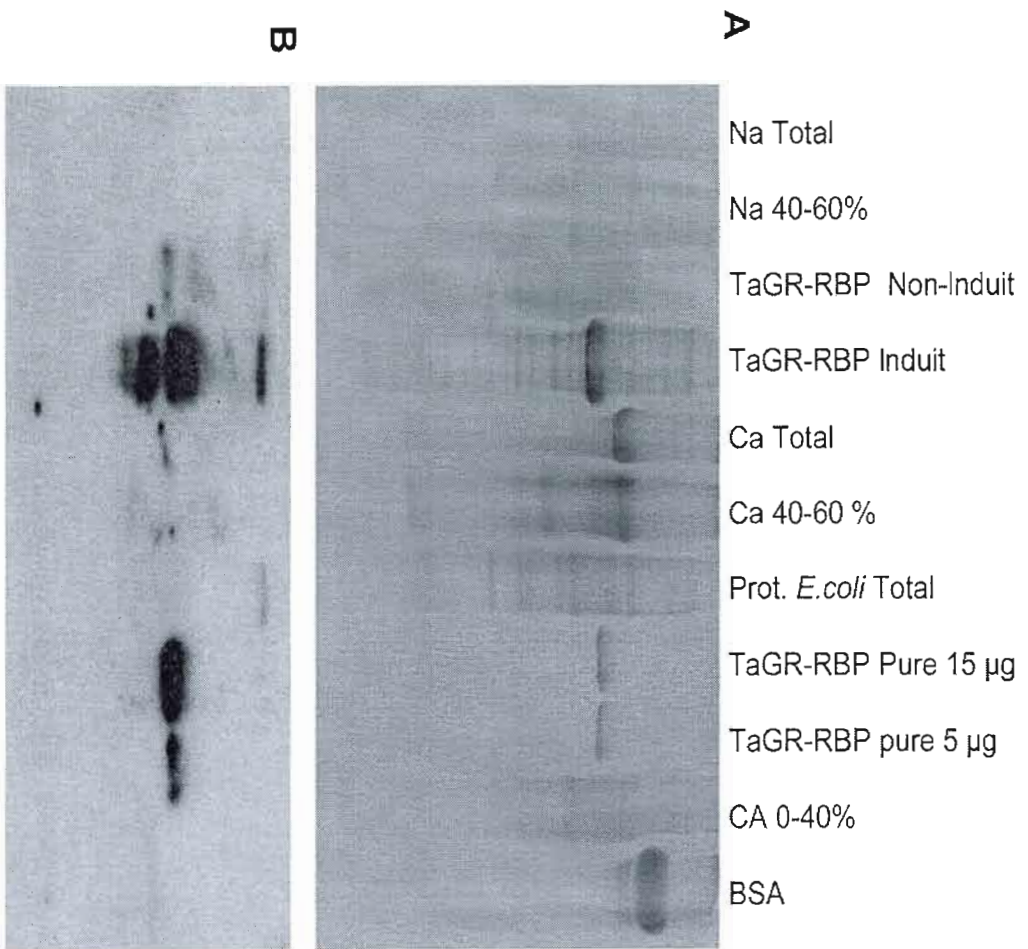
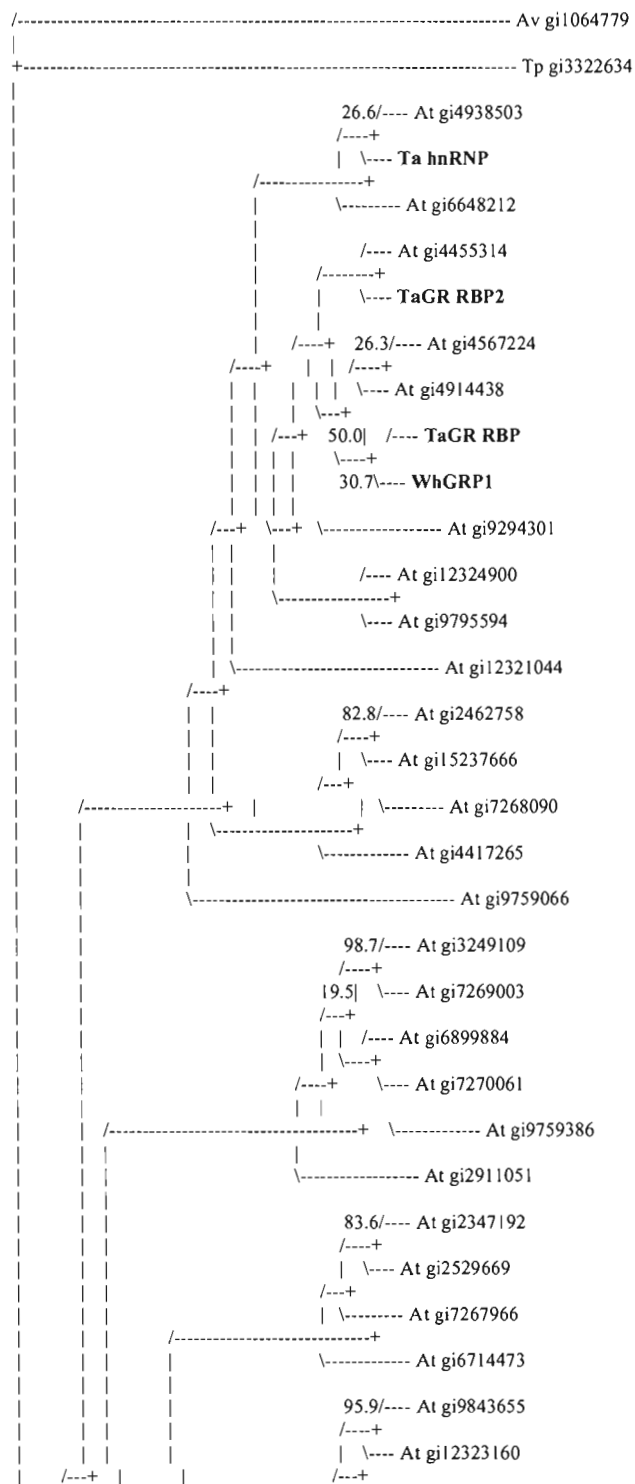
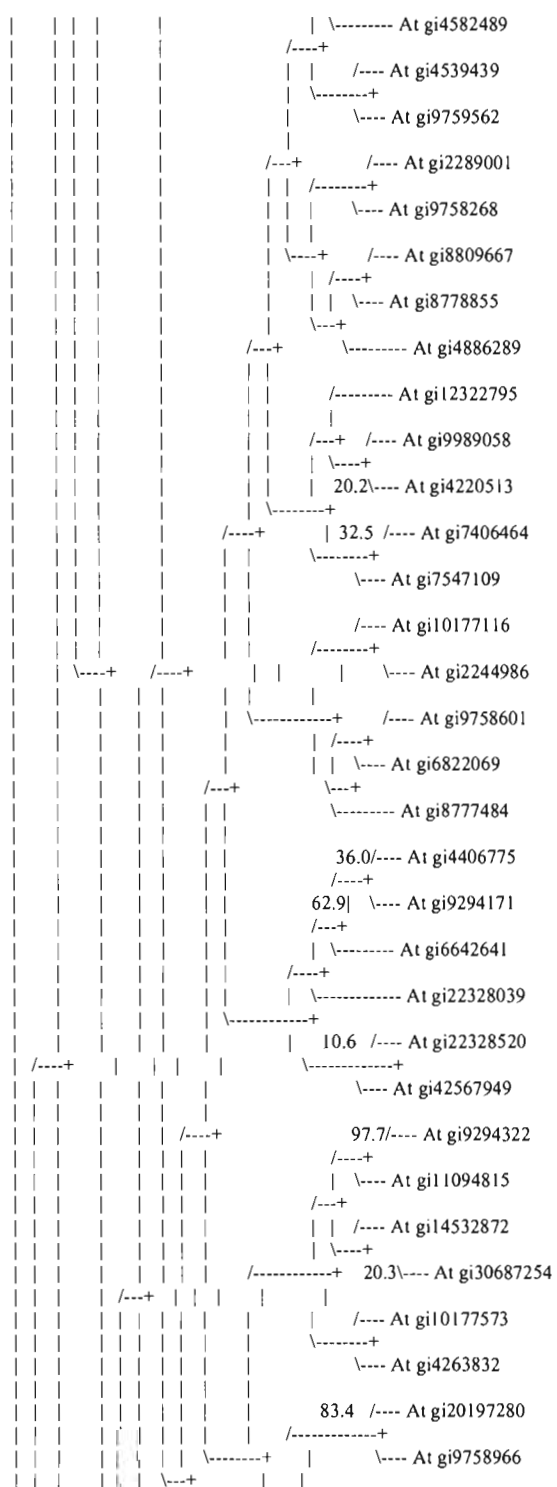


Figure 11





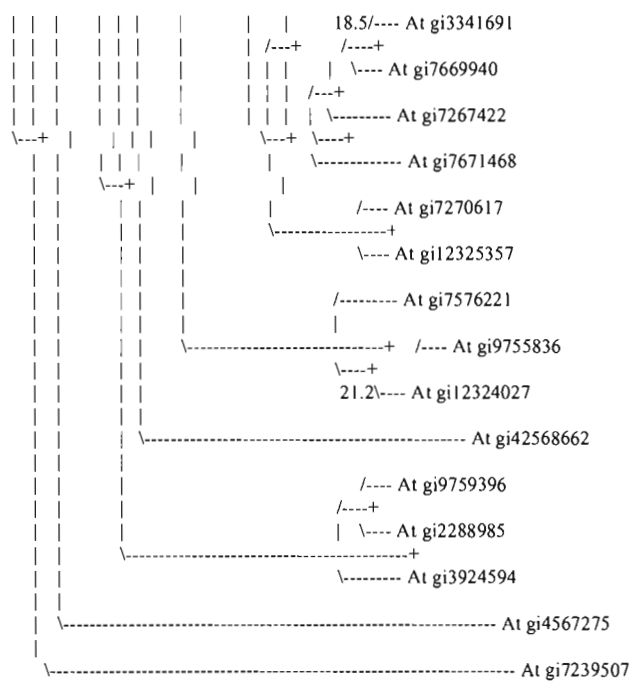


Figure 12

3.0 CONCLUSION

Nous avons lors de cette étude isolé 4 gènes de la famille des protéines liant l'ARN. Nous avons caractérisé en détail un gène dont ses homologues chez d'autres organismes étaient reconnus comme étant impliqués lors de baisses de températures. Le gène fut nommé TaGR-RBP, pour *T. aestivum* L. Glycine rich RNA binding protein. L'analyse de la séquence a démontré que cette protéine possède son domaine RRM du côté N-terminal et sa portion riche en glycine en C-terminal. Une étude d'expression par northern blot a permis de déterminer que le transcrit s'accumule chez les génotypes acclimatés comparativement aux témoins non acclimatés. L'accumulation est constante chez les différents génotypes, ce qui laisse croire que ce transcrit est induit au froid mais n'est pas directement relié à la tolérance au gel. Les transcrits de *Tagr-rbp* sont accumulés en condition de chaleur intense, de stress salin, et lors des baisses de températures. Ceci démontre que ce gène n'est pas spécifique au froid, mais qu'il est impliqué aussi dans divers stress abiotiques.

Une étude a démontré grâce à l'anticorps anti-TaGR-RBP que cette protéine s'accumule chez les variétés de blé Manitou et Norstar. L'accumulation est progressive et elle est directement proportionnelle au nombre de jours d'acclimatation. Cette accumulation de la protéine est identique chez ces 2 cultivars, ce qui nous donne la certitude que ce gène est impliqué dans la réponse au froid mais qui n'est pas essentiel à la tolérance au gel. Le gène *TaGr-rbp* a été localisé sur le chromosome 5 du génome du blé. Ce chromosome est reconnu comme porteur de plusieurs facteurs de transcription impliqués lors de la régulation de la tolérance au gel chez le blé.

D'après les analyses des séquences et les homologues de plusieurs espèces de la protéine TaGR-RBP, nous avons proposé que cette protéine possède une fonction commune entre les différents organismes. Ceci s'explique du fait que plusieurs

organismes possèdent plusieurs protéines riches en glycine liant l'ARN et que possiblement ils auraient une fonction similaire due à leurs séquences extrêmement conservés. La protéine TaGR-RBP pourrait jouer un rôle dans le métabolisme des ARN ribosomiques qui sont très actifs lors d'une baisse de température afin de pouvoir traduire les protéines nécessaires ayant un rôle direct ou indirect lors de l'acclimatation au froid. Cette hypothèse sera confirmée par des études d'interaction protéine-protéine présentement en cours. Des études de fonction ont débuté pour surexprimer le gène *TaGr-rbp* chez les plantes transgéniques afin de déterminer sa fonction.

4.0 RÉFÉRENCES

- Abler M.L., Green P.J. 1996. Control of mRNA stability in higher plants. *Plant Molecular Biology* 32:63-78.
- Albà, M.M., Pagès, M. 1998. Plant protein containing the RNA-recognition motif. *Trends in Plant Science* 3, 15-21.
- Allard, F., Houde, M., Krol, M., Ivanov, A., Huner, N.P.A, et Sarhan, F. 1998. Betaine improves freezing tolerance in wheat. *Plant Cell Physiol.* 39: 1194-1202.
- Bisaillon, M. 2001. La structure-coiffe des ARN messagers *Medicine/sciences* 17 :312.9.
- Boyle. B., 2001. Répression du gène de défense PR-10a par SEBF chez la pomme de terre. Thèse de doctorat , Université de Montréal p.15-45.
- Bohnert HJ, Nelson DE, Jensen RG. 1995. Adaptations to Environmental Stresses. *Plant Cell.*7(7),1099-1111.
- Bray EA. 1993. Molecular Responses to Water Deficit. *Plant Physiol.*103(4),1035-1040.
- Breton, G., Danyluk, J., Ouellet, F., and Sarhan, F. 2000. Biotechnological applications of plant freezing associated proteins. *Biotechnol. Annu. Rev.* 6, 59-101.
- Breton, G., 2003. Identification de nouveaux gènes associés à la tolérance au gel et à la vernalisation chez le blé (*Triticum aestivum*) et *Arabidopsis thaliana* par des approches génomiques. Thèse de doctorat , Université du Québec à Montréal.
- Browse J, Xin Z., 2001 Temperature sensing and cold acclimation. *Curr.Opin.Plant Biol.* 4(3), 241-6.
- Calkhoven CF, Ab G. 1996. Multiple steps in the regulation of transcription-factor level and activity. *Biochem J.*15;317, 329-42.
- Castonguay Y., Nadeau P., Laberge S. 1993. Freezing tolerance and alteration of translatable mRNAs in alfalfa (*Medicago sativa* L.) hardened at subzero temperatures. *Plant Cell Physiol.* 34:31-38.
- Carson JH, Cui H, Krueger W, Schlepchenko B, Brumwell C, Barbarese E. 2001. RNA trafficking in oligodendrocytes. *Results Probl Cell Differ.*34,69-81.

- Chauvin L.P., Houde M., Sarhan F. 1993. A leaf-specific gene stimulated by light during wheat acclimation to low temperature. *Plant Mol.Biol.* 23:255-265.
- Dallaire, S., Houde, M., Gagné, Y., Saini, H.S., Boileau, S., Chevrier, N., Sarhan, F. 1994. ABA and low temperature induce freezing tolerance via distinct regulatory pathways in wheat. *Plant Cell Physiol.* 35: 1-9.
- Danyluk, J., Perron, A., Houde, M., Limin, A., Fowler, B., Benhamou, N. Sarhan, F. 1998. Accumulation of an acidic dehydrin in the vicinity of the plasma membrane during cold acclimation of wheat. *Plant Cell.* 10: 623-638.
- Danyluk, J., Carpentier, E. Sarhan, F. 1996. Identification and characterization of a low temperature regulated gene encoding an actin-binding protein from wheat. *FEBS Lett.* 389: 324-7.
- Danyluk.J., 1996. Indentification et caractérisation moléculaire de gène induits au cours de l'acclimatation au froid chez le blé (*Triticum aestivum* L.), Thèse de doctorat, Université de Montréal.
- Didier DK, Klee HJ .1992. Identification of an Arabidopsis DNA- binding protein with Homology to nucleolin. *Plant Mol Biol.*18(5),977-9.
- Dragon F, Pogacic V, Filipowicz W. 2000. In vitro assembly of human H/ACA small nucleolar RNPs reveals unique features of U17 and telomerase RNAs. *Mol Cell Biol.*;20(9):3037-48.
- Dreyfuss G, Matunis MJ, Pinol-Roma S, Burd CG.1993. hnRNP proteins and the Biogenesis of mRNA.*Annu Rev Biochem* 62,289-321.
- Duchaine.T. 2001. Dynamique moléculaire et fonctions des protéines Stau 1 et Stau 2 chez les mammifères. Thèse de Doctorat Université de Montréal.
- Duncker B.P., Koops M.D., Walker V.K., Davies P.L. 1995 Low temperature persistence of type I antifreeze protein is mediated by cold specific mRNA stability. *FEBS Letters* 377:185-195.
- Dunn M A., Brown K., Lightowlers R., Hughes A.M .1996. A low-temperature-responsive gene from barley encodes a protein with single-stranded nucleic acid-binding activity which is phosphorylated in vitro. *Plant Mol. Biol.* 30: 947-959.

- Dunn M.A., Goddard N.J., Zhang L., Pearce R.S., Hughes A.M. 1994. Low temperature- responsive barley genes have different control mechanisms. *Plant Molecular Biology* 24:879:888.
- FAO. 2001. The state of food and agriculture 2001. Food and agriculture organization of the United Nations. Rome (<http://www.fao.org>).
- Fowler D.B., Sarhan F., Limin A.E., Chauvin L.P. 1996. The regulatory role of vernalization in the expression of low-temperature induced genes in wheat and rye. *Theoretical and applied genetics*.
- Fedoroff, N. 2002. RNA-Binding proteins in plants: the tip of an iceberg ? *Curr.opinion in plant biol.*5: 452-459.
- Galiba G., Quarrie S.A., Sutka J., Morgounov A., Snape J.W. 1995. RFLP mapping of the vernalization (*Vrn1*) and frost resistance (*Fr1*) genes on chromosome 5A of wheat. *Theoretical and applied Genetics* 90:1174-1179.
- Gallie D.R. 1993. Posttranscription regulation of gene expression in plants. *Ann. Rev.of Plant Physiol. and Plant Mol. Biol.* 44:77-105.
- Gautrey H, McConnell J, Hall J, Hesketh J. 2005. Polarised distribution of the RNA-binding protein Staufen in differentiated intestinal epithelial cells. *FEBS Lett.* 579(10), 2226-30.
- Gilmour SJ, Sebolt AM, Salazar MP, Everard JD, Thomashow MF. 2000. Overexpression of the *Arabidopsis* CBF3 transcriptional activator mimics multiple biochemical changes associated with cold acclimation. *Plant Physiol.* 124(4):1854-65.
- Gilmour, S.J., D.G. Zarka, E.J. Stockinger, M.P. Salazar, J.M. Houghton and M.F. Thomashow. 1998. Low temperature regulation of the *Arabidopsis* CBF family of AP2 transcriptional activators as an early step in cold-induced COR gene expression. *Plant J.* 16: 433-443.
- Grishin N.V. 2001. KH domain: one motif, two folds. *Nucleic acids res.* vol 29 no3.
- Gray GR, Chauvin LP, Sarhan F, Huner N. 1997. Cold Acclimation and Freezing Tolerance (A Complex Interaction of Light and Temperature). *Plant Physiol.* 4(2),467-474.
- Guy C.L. 1990 Cold acclimation and freezing stress tolerance: role of protein metabolism. *Ann. Rev. Plant Physiol. and Plant Mol. Biol.* 41: 187-223.

Hajela R.K., Horvath D.P., Gilmour S.J., Thomashow M.F. 1990. Molecular cloning and expression of *cor* (cold regulated) genes in *Arabidopsis thaliana*. Plant Physiol. 93,1246:1252.

Hanna-Rose W, Hansen U. 1996. Active repression mechanisms of eukaryotic Transcription repressors Trends Genet.(6):229-34.

Harmer, S., Hogenesch, J., Straume, M., Chang, H., Han,B., Zhu,T., Wang, X., Kreps, J., Kay, S. 2000., Orchestrated Transcription of key pathways in arabidopsis by the circadian clock. Science 290, 2110-2113.

Hayes P.M., Blake T., Chen T.H.H., Tragoonrung S., Chen F., Pan A., Lui B. 1993. Quantitative trait loci on barley (*Hordeum vulgare* L.) chromosome 7 associated with components of winter hardiness. Genome 36:66-71.

Hoek KS, Kidd GJ, Carson JH, Smith R.1998. hnRNP A2 selectively binds the cytoplasmic transport sequence of myelin basic protein mRNA. Biochemistry. 37(19), 7021-9.

Horton R.H, Moran A.L, Ochs S.R, Rawn D.J, Scrimgeour G.K. 1994. Principes de Biochimie, De Boeck-Wesmael Bruxelles.

Hunter,T., Karin,M. 1992. The regulation of transcription by phosphorylation. Cell 70:375-387.

Ingram J, Bartels D. 1996. The molecular basis of dehydration tolerance in plants. Annu.Rev.Plant.Physiol.Plant.Mol.Biol.47,377-403.

Jaglo-Ottosen KR, Gilmour SJ, Zarka DG, Schabenberger O, Thomashow MF. 1998. Arabidopsis CBF1 overexpression induces COR genes and enhances freezing tolerance. Science.(280),36.

Jofuku KD, den Boer BG, Van Montagu M, Okamuro JK. 1994. Control of Arabidopsis flower and seed development by the homeotic gene APETALA2. Plant Cell.(9):1211-25.

Lau JS, Baumeister P, Kim E, Roy B, Hsieh TY, Lai M, Lee AS. 2000. Heterogeneous nuclear ribonucleoproteins as regulators of gene expression through interactions with the human thymidine kinase promoter. J.Cell Biochem.79(3),395-406.

Liu Q, Kasuga M, Sakuma Y, Abe H, Miura S, Yamaguchi-Shinozaki K, Shinozaki K. 1998. Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in *Arabidopsis*. *Plant Cell*. (8):1391-406.

Luo MJ, Reed R.1999. Splicing is required for rapid and efficient mRNA export in metazoans.. *Proc Natl Acad Sci U S A*. 96(26),14937-42.

Lodish.H. Baltimore D, Darnell.J, 1993. Biologie moléculaire de la cellule. 2e éd. De Boeck, Bruxelles, Belgique.

Lorkovic L., Barta.A. 2002. Genome analysis :RNA recognition motif (RRM) and K homology (KH) domain RNA-Binding proteins from the flowering plant *Arabidopsis thaliana* *Nucleic Acids Research*. (3) 623-635.

Luscombe,N.M., Austin, S.E., Berman, H.M., Thornton, J.M. 2000. An overview of the structure of protein-DNA complexes *Genome Biol*.1: 1-37.

Magnani E, Sjolander K, Hake S. 2004. From endonucleases to transcription factors: evolution of the AP2 DNA binding domain in plants. *Plant Cell*.(9),2265-77.

Maruyama K, Sato N, Ohta N .1999. Conservation of structure and cold-regulation of RNA-binding proteins in cyanobacteria: probable convergent evolution with eukaryotic glycine-rich RNA-binding proteins *Nucleic Acids Res*.27(9), 2029-36

Maxson JM, Woodson WR.*Plant Mol Biol*. 1996. Cloning of a DNA-binding protein that interacts with the ethylene-responsive enhancer element of the carnation GST1 gene. (4):751-9.

McIntosh R.A. 1987. Gene location and gene mapping in hexaploid wheat. In: wheat and wheat improvement (Heyne, E.G., ed). *Agronomy monograph no.13* (2nd edition) p.269-287, Madison, USA.

Meshi T, Iwabuchi M.1995. Plant transcription factors. *Plant Cell Physiol*. 36(8),1405-20.

Meyer EL, Gavis ER. 2005. Staufen does double duty. *Nat Struct Mol Biol*.(4):291-2.

Monroy AF, Dhindsa RS. 1995. Low-temperature signal transduction: induction of cold acclimation-specific genes of alfalfa by calcium at 25 degrees C. *Plant Cell*.(3):321-31.

- Monroy, A.F., Sarhan, F. et Dhindsa, R.S. 1993. Cold-induced changes in freezing tolerance, protein phosphorylation, and gene expression. *Plant Physiol.* 102: 1127-35.
- Ohta M, Matsui K, Hiratsu K, Shinshi H, Ohme-Takagi M. 2001. Repression domains of class II ERF transcriptional repressors share an essential motif for active repression. *Plant Cell.*(8),1959-68.
- Ouellet.F., 1998. Caractérisation du promoteur du gène wcs120, un gène induit au cours de l'acclimatation au froid chez le blé (*Triticum aestivum* L.), Thèse de doctorat, Université de Montréal.
- Ouellet, F., Houde, M. et Sarhan, F. 1993. Purification, characterization and cDNA cloning of the 200 kDa protein induced by cold acclimation in wheat. *Plant cell Physiol.* 34: 59-65.
- Pabo C.O., Sauer R.T. 1992. Transcription factors: structural families and principles of DNA recognition. *Annu.Rev.Biochem.* 61:1053-1095.
- Perras M., Sarhan F. 1984. Energy state of spring and winter wheat during cold hardening. Soluble sugars and adenine nucleotides. *Physiol.Plant.* 60:129-132.
- Riechmann JL, Meyerowitz EM. 1998. The AP2/EREBP family of plant transcription factors.. *Biol Chem.* (379) ,633-46.
- Pearce RS. 1999 . Molecular analysis of acclimation to cold. *Plant Growth Regulation* 29, 47-76.
- Parvin, J., Sharp, P.A. 1993. DNA Topology and a minimal set of basal factors for transcription by RNA polymerase II. *Cell* 73: 533-540.
- Sarhan F., Chevrier.N. 1985. Regulation of RNA synthesis by DNA-dependant RNA polymerase and RNases during cold acclimation in winter and spring wheat. *Plant Physiol.* 78:250-255.
- Sarhan, F., and Danyluk, J. 1998. Engineering cold-tolerant crops : throwing the master switch. *Trends Plant Sci.* 3, 289-290.
- Seguin A, Laible G, Leyva A, Dixon RA, Lamb CJ. 1997. Characterization of a gene encoding a DNA-binding protein that interacts in vitro with vascular specific cis elements of the phenylalanine ammonia-lyase promoter. *Plant Mol Biol.* (53):281-91.
- Singh J., Laroche A. 1988. Freezing tolerance in plants: a biochemical overview. *Biochemistry and Cell Biology* 66:650-657.

Siomi,H., Dreyfuss,G. 1997. RNA-Binding proteins as regulators of gene expression. *Current opinion in genetics and development* 7:345-353.

Siomi, H., Choi, M., Siomi, M.C., Nussbaum, R.L.,Dreyfuss, G. 1994 . Essential role for KH domains in RNA binding : impaired RNA binding by a mutation in the KH domain of FMR1 that causes fragile X chromosome. *Cell* 77: 83-39.

Staiger,D. 2001. RNA-binding proteins and circadian rhythms in *Arabidopsis thaliana* *Philos Trans R Soc Lond B Biol Sci.*:356,1755-9.

Steponkus PL, Lynch DV. 1989. *J Bioenerg Biomembr.* Freeze/thaw-induced destabilization of the plasma membrane and the effects of cold acclimation.1 , 21-41.

Stockinger EJ, Gilmour SJ, Thomashow MF. 1997. *Arabidopsis thaliana* CBF1 encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. *Proc Natl Acad Sci U S A.* 94,1035-40.

Sutka J. 1994. Genetic control of frost tolerance in wheat *Triticum aestivum* L. *Euphytica* 77:277-282.

Suzuki M, Kuroda C, Oda E, Tsunoda S, Nakamura T, Nakajima T, Oda K. 1995. G10BP, an E1A-inducible negative regulator of Sp1, represses transcription of the rat Fibronectin gene. *Mol Cell Biol.*10, 5423-33.

Thomashow, M.F. 1999 . Role of cold-responsive genes in plant freezing tolerance. *Plant Physiol.* 118, 1-8.

Thomashow, M.F. 1999. Plant cold acclimation: Freezing tolerance genes and regulatory mechanisms. *Annu Rev Plant Physiol Plant Mol Biol.* 50:571-599.

Thomson AM, Rogers JT, Walker CE, Staton JM, Leedman PJ. 1999. Optimized RNA gel-shift and UV cross-linking assays for characterization of cytoplasmic RNA-protein interactions. *Biotechniques.* 5, 1032-9.

Tijan,R., Maniatis, T. 1994 . Transcriptional activation: a complex puzzle with few easy pieces. *Cell* 77: 5-8.

Uemura M, Joseph RA, Steponkus PL. 1995. Cold Acclimation of *Arabidopsis thaliana* Effect on Plasma Membrane Lipid Composition and Freeze-Induced Lesions .*Plant Physiol.*Sep;1091,15-30.

Williams ME., Foster R., Chua NH. 1992. Sequences flanking the hexameric G-box core CACGTG affect the specificity of protein binding. *Plant Cell*. 4 ,485-96.

Wolfraim L.A., Langis R., Tyson H., Dhindsa R.S. 1993. cDNA sequence, expression, and transcript stability of a cold acclimation-specific gene, *cas18*, of alfalfa *Medicago falcata* cells. *Plant Physiol*. 101:1275-1282.

Xin, Z., and Browse, J. 2000. Cold comfort farm: the acclimation of plants to freezing temperatures. *Plant Cell Environ*. 23, 893-902.

Zhou, B.L, Arakawa,K. Fujikawa,S.,Yoshida,S., 1994. Cold induced alteration in plasma membrane proteins that are specifically related to the development of freezing tolerance in cold-hardy winter wheat. *Plant Cell Physiol*.35:175-182.

5.0 ANNEXE

Tableau I : Liste de quelques domaines auxiliaires se trouvant chez la famille de protéines liant l'ARN

Domaines auxiliaires	fonctions désignés
Dipeptides SR	Inconnue
C2HC jointure de Zinc	impliqué dans la liaison simple brin de nucléotide
C3H doit de zinc	impliqué dans la liaison simple brin de nucléotide
portion acide	facilite la liaison a l'ARN
Riche en Q	Inconnue
Riche en G	impliqué lors de différentes stress abiotiques
Riche en GY	Inconnue
KRDE riche	Inconnue
Ancre RD	Inconnue
Domaine NTF	Impliqué lors de l'épissage de transcrit
Riche en PQS	Inconnue
Sequence signal au chloroplaste	transport aux chloroplastes pour stabilisation

Tableau II : Liste des numéros d'accension d'*Arabidopsis thaliana* ayant un domaine de liaison à l'ARN RRM et ceux de leurs homologues chez le riz avec leurs valeurs d'homologie (E value).

Numéro d'accension <i>Arabidopsis thaliana</i>	Définition du gène trouvé par BLAST chez le Riz	"E" Value
AAB63640.	gi 46389987 dbj BAD16229.1 putative poly A)-binding protein	3.00E-86
CAA17561.	gi 46389987 dbj BAD16229.1 putative poly A)-binding protein	0.0
CAA72907.	gi 46389987 dbj BAD16229.1 putative poly A)-binding protein	1.00E-162
AAB87097.	gi 46389987 dbj BAD16229.1 putative poly A)-binding protein	1.00E-154
AAF43230.	gi 46389987 dbj BAD16229.1 putative poly A)-binding protein	1.00E-167
AAD20142.	gi 46389987 dbj BAD16229.1 putative poly A)-binding protein	1.00E-94
AAG13056.	gi 46389987 dbj BAD16229.1 putative poly A)-binding protein	0.0
BAB11475..	gi 50725435 dbj BAD32907.1 putative polyadenylate-binding protein	3.00E-18
NP_175125.	gi 50926280 ref XP_473087.1 OSJNBa0014K14.18	3.00E-15
AAG12523.	gi 46389987 dbj BAD16229.1 putative poly A)-binding protein	2.00E-86
CAB89408.	gi 46805689 dbj BAD17090.1 putative poly A) binding protein	9.00E-47
AAK61429.	gi 24413950 dbj BAC22202.1	2.7

	hypothetical protein	
BAA97374.	gi 51535367 dbj BAD37238.1	8.00E-44
	putative poly A) binding protein II	
AAC24092.	gi 33146854 dbj BAC79849.1	2.00E-68
	putative pre-mRNA splicing factor SF2	
CAB80736.	gi 33146854 dbj BAC79849.1	2.00E-58
	putative pre-mRNA splicing factor SF2	
NP_190512.	gi 33146854 dbj BAC79849.1	3.00E-67
	putative pre-mRNA splicing factor SF2	
CAB80736	gi 33146854 dbj BAC79849.1	2.00E-58
	putative pre-mRNA splicing factor SF2	
CAB71893.	gi 50921405 ref XP_471063.1	3.00E-75
	OSJNBa0020P07.8	
CAB81360.	gi 41053015 dbj BAD07946.1	2.00E-65
	putative arginine/serine-rich splicing factor RSp41	
AAK43986.	gi 41053015 dbj BAD07946.1	9.00E-74
	putative arginine/serine-rich splicing factor RSp41	
AAF87159.	gi 51091721 dbj BAD36521.1	4.00E-35
	putative splicing factor, arginine/serine-rich 7	
CAB79876.	gi 51091721 dbj BAD36521.1	2.00E-37
	putative splicing factor, arginine/serine-rich 7	
AAD23894.1	gi 51091721 dbj BAD36521.1	3.00E-37
	putative splicing factor, arginine/serine-rich 7	
CAB67657.	gi 54290784 dbj BAD61423.1	1.00E-59
	putative splicing factor	
AAC98054.	gi 54290784 dbj BAD61423.1	1.00E-49
	putative splicing factor	
CAC03601.	gi 33146902 dbj BAC79901.1	2.00E-27
	putative SC35-like splicing factor SCL30a	

CAB75904.	gi 32352198 dbj BAC78592.1	4.00E-31
	pre-mRNA splicing factor	
AAK44083.	gi 31249706 gb AAP46199.1	1.00E-41
	putative splicing factor	
AAG51556.	gi 31249706 gb AAP46199.1	4.00E-41
	putative splicing factor	
BAB09851.1	gi 33146621 dbj BAC79909.1	8.00E-48
	putative splicing factor, arginine/serine-rich 2 Splicing factor SC35)	
AAG10821.	gi 57900284 dbj BAD87117.1	6.00E-35
	putative Ser/Arg-related nuclear matrix protein	
AAC62852.	gi 34897622 ref NP_910157.1	3.00E-82
	putative small nuclear ribonucleoprotein U1A [Oryza sativa]	
CAB62436.	gi 31429881 gb AAP51870.1	2.00E-76
	putative U1 small nuclear ribonucleoprotein	
AAC16931.	gi 34897622 ref NP_910157.1	1.00E-64
	putative small nuclear ribonucleoprotein U1A [Oryza sativa]	
AAF82223.	gi 34897622 ref NP_910157.1	2.00E-54
	putative small nuclear ribonucleoprotein U1A [Oryza sativa]	
CAB80335.1	gi 29367529 gb AAO72620.1	1.00E-140
	putative U2 snRNP auxiliary factor	
AAG51641.1	gi 29367529 gb AAO72620.1	1.00E-133
	putative U2 snRNP auxiliary factor	
AAD12222.1	gi 31432465 gb AAP54095.1	1.00E-119
	putative spliceosomal protein	
AAD15475.1	gi 31432465 gb AAP54095.1	3.00E-69
	putative spliceosomal protein	
BAA97393.1	gi 50923815 ref XP_472268.1	1.00E-148
	OSJNBa0044M19.2	
AAB64330.1	gi 50725348 dbj BAD34420.1	1.00E-69
	ribonucleoprotein antigen-like	

BAB10638.1	gi 15239067 ref NP_199096.1	1.00E-120
	U2 snRNP auxiliary factor small subunit, putative [Arabidopsis thaliana]	
AAD46002.1	gi 12744991 gb AAK06875.1 AF344324_1	5.00E-92
	putative U2 snRNP auxiliary factor [Arabidopsis thaliana]	
AAD32882.1	gi 46390906 dbj BAD16421.1	2.00E-38
	U2 snRNP auxiliary factor-like protein	
AAD29697.1	gi 30692291 gb AAP33448.1	4.00E-77
	cap-binding protein CBP20	
BAB10037.1	gi 32129324 gb AAP73851.1	1.00E-54
	putative RNA binding protein	
BAB08767.1	gi 12583795 gb AAG59647.1 AC084319_5	1.00E-41
	conserved hypothetical protein protein [Oryza sativa]	
AAD30259.1	gi 40253807 dbj BAD05744.1	1.00E-107
	putative RNA Binding Protein 45	
AAG13046.1	gi 51535291 dbj BAD38554.1	9.00E-98
	putative nucleic acid binding protein	
BAB02953.1	gi 51535291 dbj BAD38554.1	4.00E-97
	putative nucleic acid binding protein	
AAD46038.1	gi 51535291 dbj BAD38554.1	1.00E-102
	putative nucleic acid binding protein	
AAD46037.1	gi 51535291 dbj BAD38554.1	1.00E-102
	putative nucleic acid binding protein	
AAD25780.1	gi 23237933 dbj BAC16506.1	1.00E-143
	putative oligouridylate binding protein	
AAF79492.1	gi 23237933 dbj BAC16506.1	1.00E-133
	putative oligouridylate binding protein	
BAB02974.1	gi 23237933 dbj BAC16506.1	1.00E-134
	putative oligouridylate binding protein	
CAC00749.1	gi 34903096 ref NP_912895.1	5.00E-86

	unnamed protein product	
AAD12005.1	gi 34903096 ref NP_912895.1	1.00E-76
	unnamed protein product	
BAA97064.1	gi 54290288 dbj BAD61233.1	3.00E-75
	heterogeneous nuclear ribonucleoprotein A2/B1-like	
AAD25815.2	gi 34903096 ref NP_912895.1	6.00E-32
	unnamed protein product	
AAD25814.2	gi 34903096 ref NP_912895.1	2.00E-25
	unnamed protein product	
AAC16468.2	gi 53793441 dbj BAD53164.1	7.00E-29
	RNA-binding protein-like	
BAB09544.1	gi 32129324 gb AAP73851.1	8.00E-24
	putative RNA binding protein	
CAB88326.1	gi 50924107 ref XP_472414.1	2.00E-17
	OJ000126_13.13	
CAB78306.1	gi 50725522 dbj BAD32991.1	2.00E-19
	RNA recognition motif (RRM)-containing protein-like	
AAG51218.1	gi 34907334 ref NP_915014.1	3.00E-41
	P0698A10.29	
AAD20390.1	gi 50881454 gb AAT85299.1	7.00E-17
	glycine-rich RNA-binding protein, putative	
AAC23648.2	gi 50940959 ref XP_480007.1	3.00E-22
	putative RNA recognition motif (RRM)-containing protein	
AAF21210.1	gi 7439986 pir T03586	4.00E-23
	glycine-rich RNA-binding protein 2 - rice	
BAB09686.1	gi 50924107 ref XP_472414.1	2.00E-24
	OJ000126_13.13	
BAB09337.1	gi 50940959 ref XP_480007.1	8.00E-39
	putative RNA recognition motif (RRM)-containing protein	

BAB08354.1	gi 50924107 ref XP_472414.1	4.00E-17
	OJ000126_13.13	
AAB88654.1	gi 52077025 dbj BAD46058.1	2.00E-29
	RNA-binding protein-like	
BAB09540.1	gi 51090851 dbj BAD35379.1	5.00E-24
	putative SEB4D	
CAB77597.1	gi 27476063 gb AAO16994.1	2.00E-37
	Hypothetical protein	
AAF27001.1	gi 34907334 ref NP_915014.1	5.00E-22
	P0698A10.29	
AAD30604.1	gi 51090851 dbj BAD35379.1	6.00E-54
	putative SEB4D	
AAG51948.1	gi 51090851 dbj BAD35379.1	3.00E-58
	putative SEB4D	
AAF87259.1	gi 51535080 dbj BAD37669.1	1.00E-33
	putative RRM-containing protein	
AAF71806.1	gi 34907334 ref NP_915014.1	2.00E-34
	P0698A10.29	
AAC33496.1	gi 34907334 ref NP_915014.1	4.00E-45
	P0698A10.29	
AAK32921.1	gi 34907334 ref NP_915014.1	8.00E-46
	P0698A10.29	
AAF20222.1	gi 48716424 dbj BAD23032.1	5.00E-17
	putative Ras-GTPase-activating protein binding protein 1	
AAD20086.1	gi 48716424 dbj BAD23032.1	4.00E-51
	putative Ras-GTPase-activating protein binding protein 1	
AAF27060.1	gi 50923623 ref XP_472172.1	4.00E-40
	OSJNBa0069D17.2	
AAF81299.1	gi 48716424 dbj BAD23032.1	8.00E-45
	putative Ras-GTPase-activating protein binding protein 1	

BAB09056.1	gi 50905045 ref XP_464011.1	9.00E-66
	putative Ras-GTPase activating protein SH3 domain-binding protein 2	
BAB10647.1	gi 50923623 ref XP_472172.1	9.00E-77
	OSJNBa0069D17.2	
BAB10698.1	gi 48716424 dbj BAD23032.1	1.00E-55
	putative Ras-GTPase-activating protein binding protein 1	
BAB02073.1	gi 48716424 dbj BAD23032.1	6.00E-73
	putative Ras-GTPase-activating protein binding protein 1	
AAF20221.1	gi 48716424 dbj BAD23032.1	3.00E-31
	putative Ras-GTPase-activating protein binding protein 1	
CAB10407.1	gi 50261763 gb AAT72462.1	1.00E-103
	FCA gamma protein	
AAB63831.1	gi 50261763 gb AAT72462.1	9.00E-48
	FCA gamma protein	
AAB64314.1	gi 50725369 dbj BAD34441.1	1.00E-110
	unknown protein	
AAG52616.1	gi 50540745 gb AAT77901.1	3.00E-34
	putative RNA binding motif protein	
NP_200803.3	gi 38175495 dbj BAD01191.1	7.00E-33
	putative transcriptional coactivator DIP2)	
AAG51767.1	gi 52077258 dbj BAD46300.1	2.00E-31
	putative DIP2 protein	
CAB85990.1	gi 38175495 dbj BAD01191.1	4.00E-35
	putative transcriptional coactivator DIP2)	
NP_198588.2	gi 38175495 dbj BAD01191.1	1.00E-41
	putative transcriptional coactivator DIP2)	
AAG51445.1	gi 47497762 dbj BAD19862.1	1.00E-108
	putative initiation factor 3g	
BAB10805.1	gi 47497762 dbj BAD19862.1	6.00E-94

	putative initiation factor 3g	
BAA97166.1	gi 60498747 dbj BAC19887.2	4.00E-25
	Ribosomal protein S19	
CAB41335.1	gi 49388840 dbj BAD26030.1	1.00E-52
	putative plastid-specific ribosomal protein 2 precursor	
AAF43233.1	gi 62733690 gb AAX95801.1	9.00E-85
	RNA recognition motif. a.k.a. RRM, RBD, or RNP domain), putative	
CAB40027.1	gi 50725622 dbj BAD33089.1	1.00E-107
	putative RNA-binding protein RBP37	
AAF25974.1	gi 50725622 dbj BAD33089.1	1.00E-110
	putative RNA-binding protein RBP37	
CAB66418.1	gi 50725622 dbj BAD33089.1	4.00E-98
	putative RNA-binding protein RBP37	
BAB01337.1	gi 56201519 dbj BAD73038.1	8.00E-97
	RNA-binding protein -like	
BAB08327.1	gi 31433615 gb AAP55107.1	7.00E-46
	putative NAM no apical meristem) protein	
AAG51971.1	gi 56201519 dbj BAD73038.1	5.00E-85
	RNA-binding protein -like	
AAG50640.1	gi 31432618 gb AAP54226.1	1.00E-44
	putative RNA-binding protein	
CAB80362.1	gi 57899379 dbj BAD88026.1	5.00E-83
	RNA-binding like protein	
BAB08520.1	gi 42408771 dbj BAD10006.1	2.00E-66
	putative ribonucleoprotein	
NP_173208.1	gi 46806499 dbj BAD17623.1	3.00E-59
	putative heterogeneous nuclearribonucleoprotein A2	
NP_201200.2	gi 50929915 ref XP_474485.1	0.0
	OSJNBb0020J19.17	

CAB87285.1	gi 52076187 dbj BAD46727.1	1.00E-166
	putative AML1	
AAD21720.2	gi 45735834 dbj BAD12869.1	1.00E-154
	putative meiosis protein mei2	
AAG51742.1	gi 45735834 dbj BAD12869.1	0.0
	putative meiosis protein mei2	
CAB78814.1	gi 45735834 dbj BAD12869.1	1.00E-157
	putative meiosis protein mei2	
BAB11164.1	gi 50949165 ref XP_493940.1	1.00E-115
	ESTs AU082563 S20379), D15187 C0226), AU082476 C0226),	
AAF79557.1	gi 50949165 ref XP_493940.1	1.00E-171
	putative RNA-binding protein. AC004561)	
AAC95181.1	gi 50949165 ref XP_493940.1	1.00E-172
	putative RNA-binding protein. AC004561)	
CAB61958.1	gi 15042832 gb AAK82455.1 AC091247_22	9.00E-81
	putative RNA binding protein [Oryza sativa]	
CAB10209.1	gi 31432618 gb AAP54226.1	1.00E-64
	putative RNA-binding protein	
AAB80680.1	gi 31432618 gb AAP54226.1	3.00E-66
	putative RNA-binding protein	
BAB08572.1	gi 38175737 dbj BAC55617.2	5.00E-97
	putative heterogeneous nuclear ribonucleoprotein A1	
CAB43861.1	gi 38175737 dbj BAC55617.2	1.00E-101
	putative heterogeneous nuclear ribonucleoprotein A1	
BAB09088.1	gi 38175737 dbj BAC55617.2	5.00E-98
	putative heterogeneous nuclear ribonucleoprotein A1	
AAF21191.1	gi 38175737 dbj BAC55617.2	2.00E-98
	putative heterogeneous nuclear ribonucleoprotein A1	
BAB10406.1	gi 50928389 ref XP_473722.1	1.00E-04
	OSJNBa0086O06.22	

BAB02497.1	gi 62733154 gb AAX95271.1	0.028
	RNA recognition motif. a.k.a. RRM, RBD, or RNP domain), putative	
BAB08421.1	gi 42407521 dbj BAD10638.1	1.00E-148
	putative polypyrimidine tract-binding protein homolog	
AAF26159.1	gi 18921326 gb AAL82531.1 AC084766_17	1.00E-163
	putative polypyrimidine tract-binding protein [Oryza sativa]	
NP_175010.2	gi 50931555 ref XP_475305.1	1.00E-151
	putative polypyrimidine tract-binding protein	
AAC79095.1	gi 34907248 ref NP_914971.1	1.00E-139
	putative ribonucleoprotein	
AAF86538.1	gi 34907248 ref NP_914971.1	1.00E-129
	putative ribonucleoprotein	
AAD14476.1	gi 47847880 dbj BAD21673.1	2.00E-72
	putative RNA-binding protein RNP1 precursor	
AAF26472.1	gi 50921411 ref XP_471066.1	1.00E-170
	OSJNBa0020P07.11	
CAB67653.1	gi 22296371 dbj BAC10140.1	3.00E-35
	putative 29 kDa ribonucleoprotein A, chloroplast precursor	
AAC98043.1	gi 18921322 gb AAL82527.1 AC084766_13	2.00E-74
	putative ribonucleoprotein [Oryza sativa]	
CAA22986.1	gi 52076131 dbj BAD46644.1	4.00E-81
	putative nucleic acid-binding protein	
BAB09396.1	gi 52076131 dbj BAD46644.1	7.00E-76
	putative nucleic acid-binding protein	
CAB43448.1	gi 34394882 dbj BAC84331.1	2.00E-59
	putative RNA-binding protein	
AAC36180.1	gi 50928389 ref XP_473722.1	2.00E-53
	OSJNBa0086O06.22	
AAD22311.1	gi 50881454 gb AAT85299.1	4.00E-27

	glycine-rich RNA-binding protein, putative	
CAB36849.1	gi 7439986 pir T03586	4.00E-30
	glycine-rich RNA-binding protein 2 - rice	
BAB10366.1	gi 19881711 gb AAM01112.1 AC098682_16	1.00E-27
	Putative RNA-binding protein [Oryza sativa]	
BAB03001.1	gi 7439986 pir T03586	2.00E-32
	glycine-rich RNA-binding protein 2 - rice	
AAG52402.1	gi 19881711 gb AAM01112.1 AC098682_16	5.00E-24
	Putative RNA-binding protein [Oryza sativa]	
AAF98412.1	gi 7439986 pir T03586	5.00E-17
	glycine-rich RNA-binding protein 2 - rice	
AAD23639.1	gi 50881454 gb AAT85299.1	4.00E-36
	glycine-rich RNA-binding protein, putative	
CAB43641.1	gi 7439987 pir T04346	3.00E-32
	glycine-rich RNA-binding protein - rice	
BAB02203.1	gi 40714682 gb AAR88588.1	1.00E-41
	putative RNA binding protein	
AAB71977.1	gi 33146923 dbj BAC79944.1	7.00E-42
	glycine-rich RNA-binding protein-like	
NP_196048.1	gi 33146923 dbj BAC79944.1	2.00E-48
	glycine-rich RNA-binding protein-like	
AAG51392.1	gi 50725127 dbj BAD33744.1	1.00E-51
	putative MADP-1 protein	
CAB78428.1	gi 31431114 gb AAP52936.1	6.00E-14
	putative RNA-binding protein	
NP_200687.1	gi 31432892 gb AAP54468.1	2.00E-74
	unknown protein	
AAG51976.1	gi 52076982 dbj BAD45991.1	1.00E-156
	peptidylprolyl isomerase-like	

BAB09234.1	gi 42408488 dbj BAD09668.1	2.00E-39
	putative cleavage and polyadenylation specific factor	
AAD31072.1	gi 52077336 dbj BAD46377.1	1.00E-35
	RNA-binding protein-like	
CAB70997.1	gi 49387752 dbj BAD26240.1	1.00E-127
	putative RNA-binding protein 10	
NP_188729.1	gi 46390469 dbj BAD15930.1	7.00E-88
	RNA recognition motif RRM)-containing protein-like	
AAD50053.1	gi 50725051 dbj BAD33184.1	4.00E-74
	RNA-binding protein-like	
AAF79854.1	gi 45593096 gb AAS68137.1	4.00E-19
	homeodomain leucine zipper protein 16	
CAB78307.1	gi 50725522 dbj BAD32991.1	5.00E-85
	RNA recognition motif RRM)-containing protein-like	
CAB80892.1	gi 62733051 gb AAX95168.1	1.00E-119
	RNA recognition motif. a.k.a. RRM, RBD, or RNP domain), putative	
CAB89227.1	gi 29371361 gb AAO72701.1	2.00E-99
	putative RNA-binding protein	
AAC64224.2	gi 31432309 gb AAP53959.1	6.00E-97
	putative splicing factor-like protein	
BAB09409.1	gi 31432309 gb AAP53959.1	1.00E-118
	putative splicing factor-like protein	
CAB96835.1	gi 50252084 dbj BAD28014.1	0.0
	putative U2-associated SR140 protein	
AAK64118.1	gi 50252084 dbj BAD28014.1	0.0
	putative U2-associated SR140 protein	
NP_197496.1	gi 18103931 emb CAC83314.1	1.00E-12
	glycine rich RNA binding protein [Oryza sativa indica cultivar-group]]	
CAC01867.1	gi 50925969 ref XP_473032.1	1.00E-119
	OSJNBa0067K08.9	

CAB40378.1	gi 50928659 ref XP_473857.1	1.00E-131
	OSJNBa008510.19	
AAG30979.1	gi 50940959 ref XP_480007.1	1.00E-18
	putative RNA recognition motif (RRM)-containing protein	
NP_197404.2	gi 7439986 pir T03586	3.00E-10
	glycine-rich RNA-binding protein 2 - rice	
NP_192643.2	gi 50946971 ref XP_483013.1	2.00E-38
	putative RNA-binding protein cp33	
AAC27473.1	gi 50927259 ref XP_473377.1	3.00E-29
	OSJNBb0103108.18	
CAC05503.1	gi 50949165 ref XP_493940.1	1.00E-23
	putative RNA-binding protein. AC004561)	
AAG28908.1	gi 20804887 dbj BAB92568.1	4.00E-50
	putative terminal ear1	
NP_189242.1	gi 20804887 dbj BAB92568.1	6.00E-97
	putative terminal ear1	
NP_200655.1	gi 34896094 ref NP_909391.1	6.00E-47
	B1189A09.32	
CAB87911.1	gi 21742315 emb CAD40772.1	1.00E-137
	OSJNBb0039F02.3	
	gi 52076599 dbj BAD45501.1	1.00E-101
	nuclear protein -like	
NP_198185.1	gi 29371361 gb AAO72701.1	3.00E-30
	putative RNA-binding protein	
AAF63781.1	gi 42409101 dbj BAD10352.1	2.00E-74
	putative RNA recognition motif (RRM)-containing protein	
AAC24054.1	gi 50945765 ref XP_482410.1	1.00E-131
	putative RNA-binding region RNP-1 and Splicing factor PWI family member protein	
BAB02219.1	gi 51449869 gb AAU01907.1	2.00E-58

	putative nucleolin [Oryza sativa indica cultivar-group]]	
AAG29744.1	gi 50928791 ref XP_473923.1	1.00E-51
	OSJNBa0058K23.21	
CAB82111.1	gi 51449869 gb AAU01907.1	1.00E-10
	putative nucleolin [Oryza sativa indica cultivar-group]]	
AAD23688.1	gi 29150367 gb AAO72376.1	1.00E-170
	putative RNA binding ribonucleoprotein	
CAB85566.1	gi 37806189 dbj BAC99692.1	1.00E-64
	putative RNA binding protein	
BAA97218.1	gi 50920179 ref XP_470450.1	6.00E-79
	unknown protein	